



US 20150089521A1

(19) **United States**(12) **Patent Application Publication**  
**Orlowski**(10) **Pub. No.: US 2015/0089521 A1**(43) **Pub. Date: Mar. 26, 2015**(54) **DATA TRANSLATION FOR VIDEO-VIEWING  
ACTIVITY**(71) Applicant: **Comcast Cable Communications,  
LLC, Philadelphia, PA (US)**(72) Inventor: **Robert Alan Orlowski, Centennial, CO  
(US)**(73) Assignee: **Comcast Cable Communications,  
LLC, Philadelphia, PA (US)**(21) Appl. No.: **14/037,352**(22) Filed: **Sep. 25, 2013****Publication Classification**(51) **Int. Cl.**  
**H04N 21/442** (2006.01)(52) **U.S. Cl.**CPC ..... **H04N 21/44222** (2013.01)USPC ..... **725/14**

(57)

**ABSTRACT**

A computer-implemented method of using Linear, DVR, and VOD video viewing activity data as input to a data translation processor which prepares that viewing activity for more efficient downstream processing by translating detailed values to aggregated values according to analyst defined translation rules in preparation for ingestion by a MapReduce Framework with the result that the MapReduce Framework needs to process less data in order to create analytical studies of second-by-second viewing activity for program, channel, house, device, viewer, demographic, and geographic attributes. The source data may be extracted from a database defined according to the Cable Television Laboratories, Inc. Media Measurement Data Model defined in "Audience Data Measurement Specification" as "OpenCable™ Specifications, Audience Measurement, Audience Measurement Data Specification" document OC-SP-AMD-101-130502 or any similar format. An analyst can use Hadoop to run more studies in less time with less hardware thus gaining greater insights into viewing activity at lower cost.

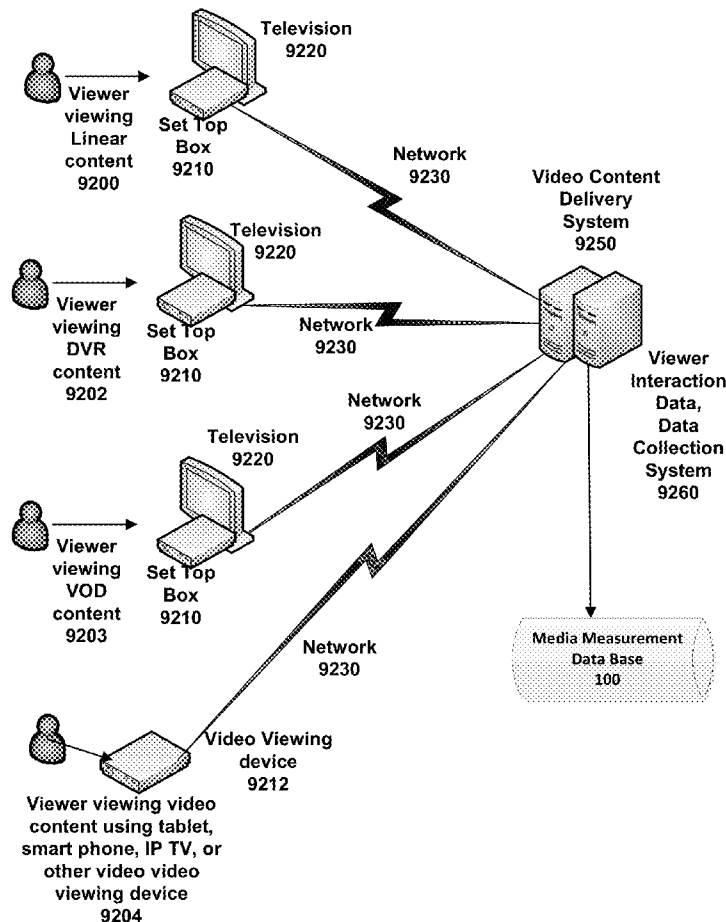


FIG. 1

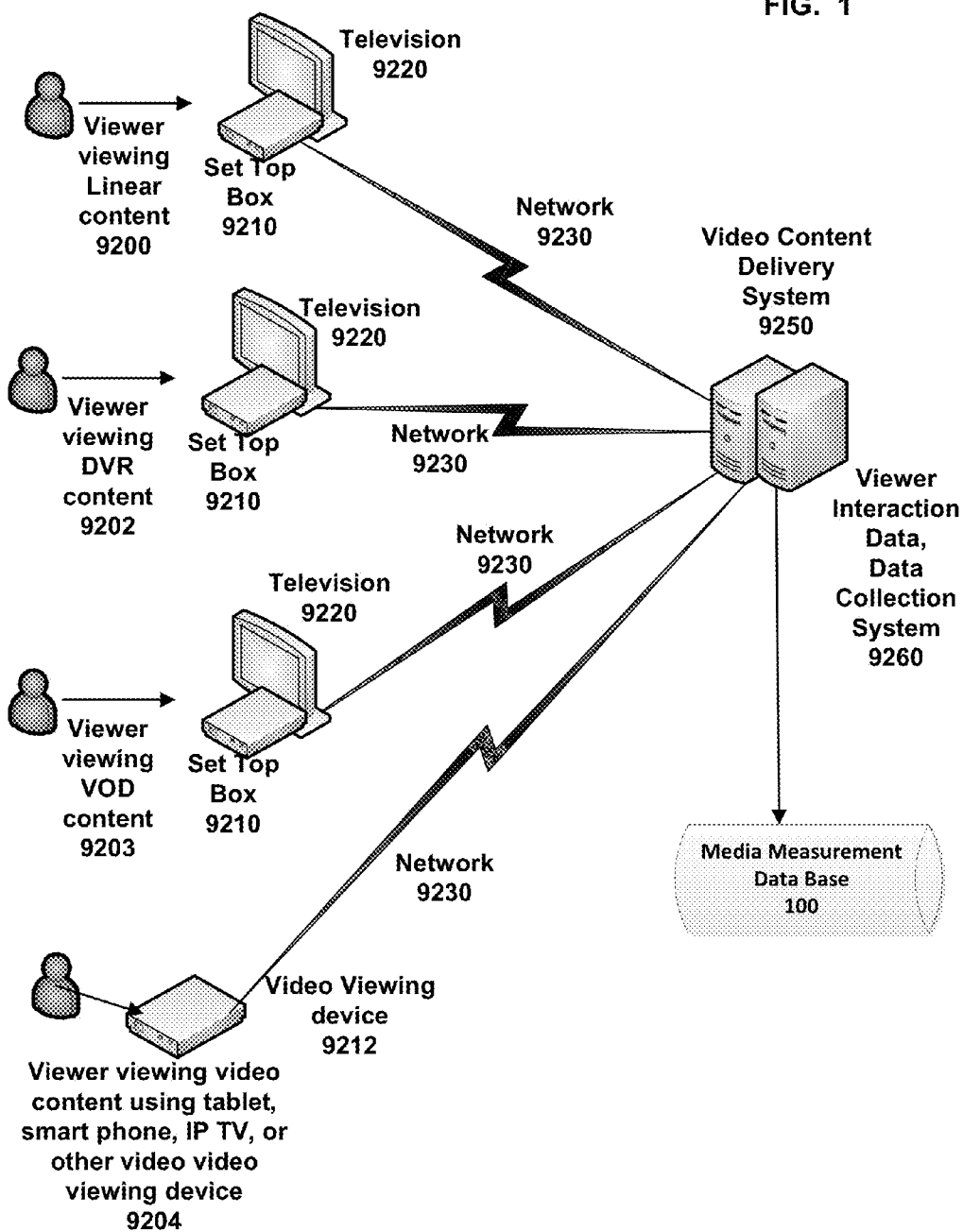


FIG. 2

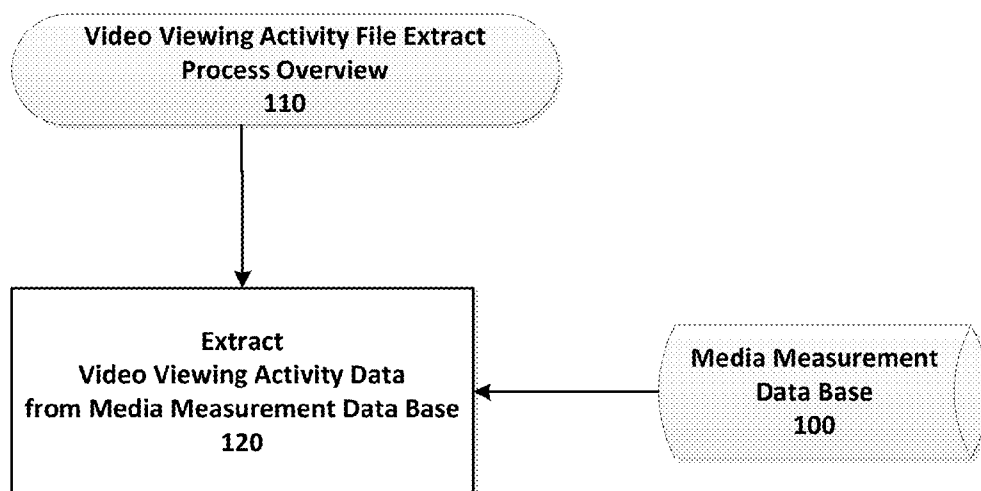


FIG. 5

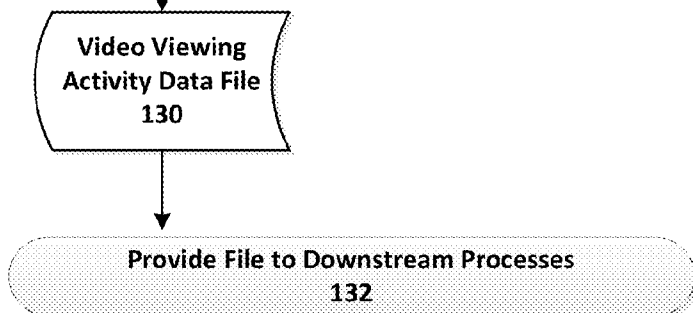


FIG. 3

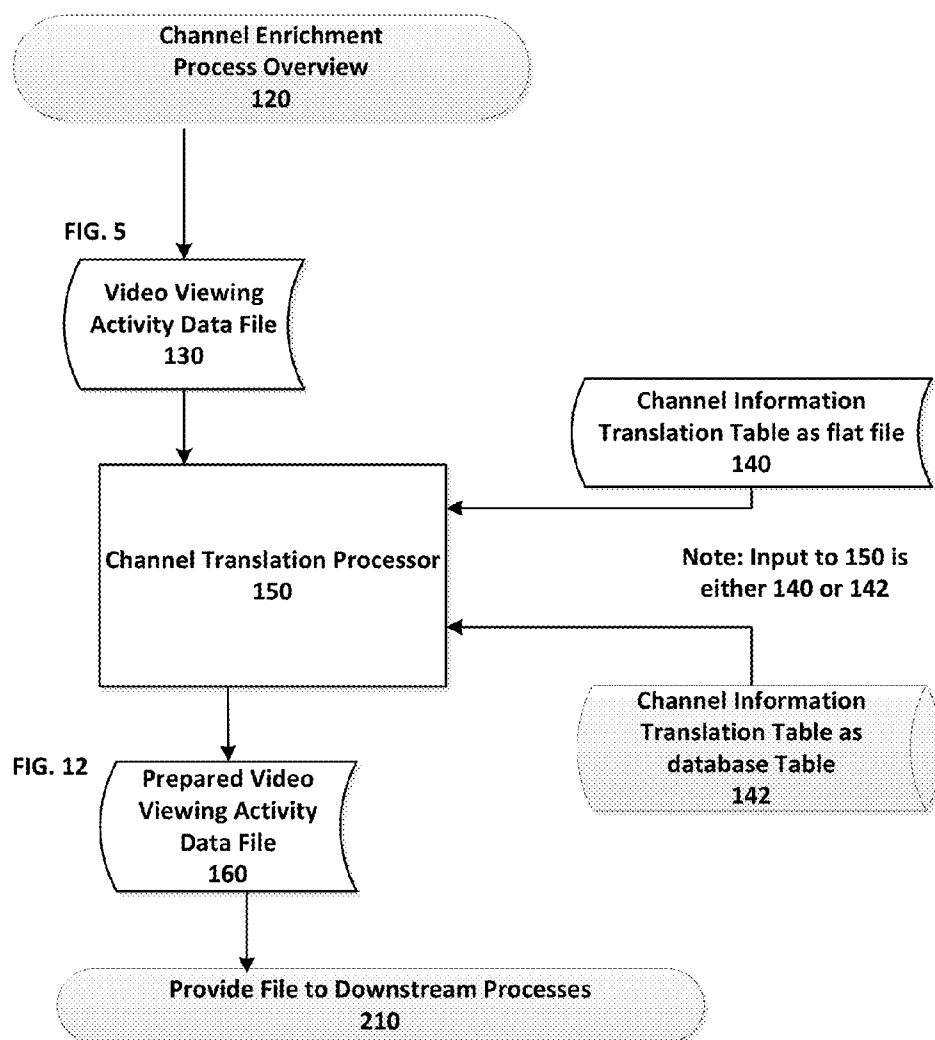


FIG. 4

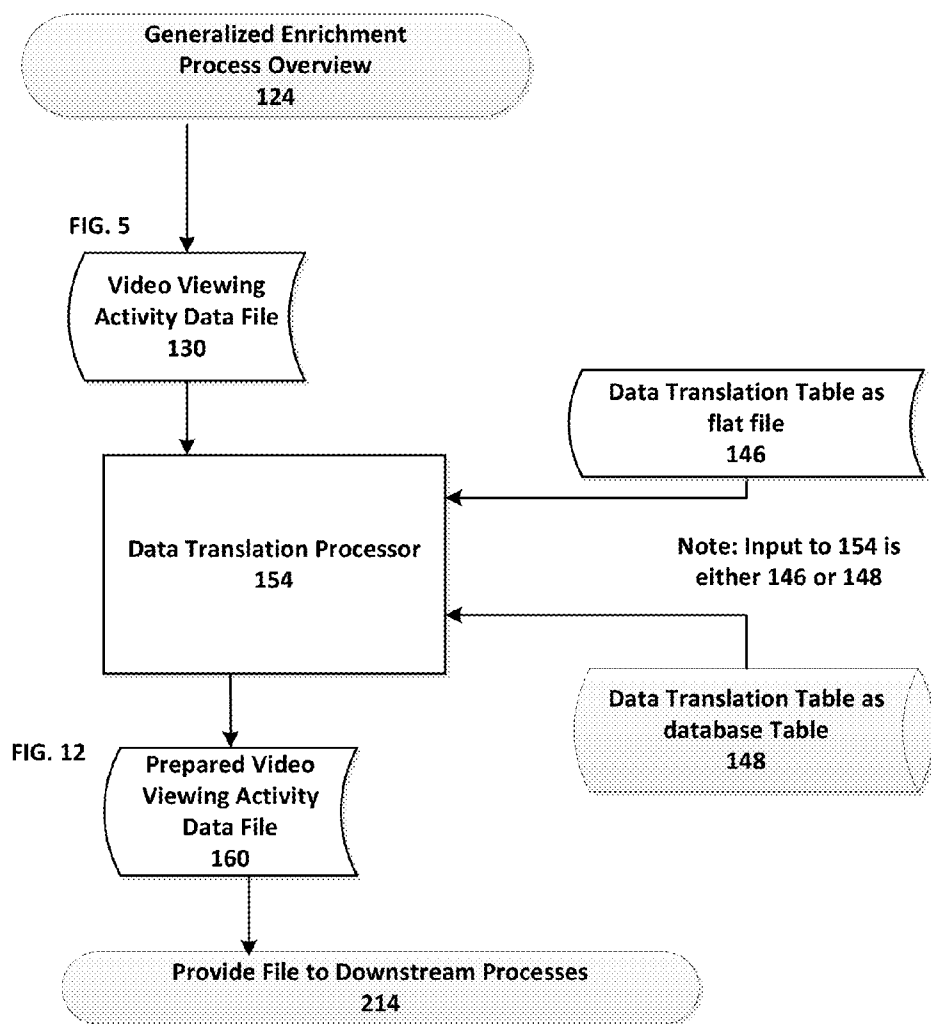


FIG. 5

## Summary Information

Specification Name:	Video Viewing Activity Data File <b>130</b>
Record description:	<p>The Video Viewing Activity Data contains video viewing activity data from a Media Measurement Data Base or a similar input source. This video viewing activity data can be enriched with additional or substitutionary fields prior to sending the file onward to the next step in aggregating the video viewing activity. The next step may be exploding the viewing activity into individual Video Viewing Detail records which can then be ingested by a MapReduce Aggregation Engine as described in my prior applications.</p> <p>The video viewing activity can be sourced from Linear Tuning Activity (<b>LTA</b>), or sessionized Linear Viewing Activity (<b>LVA</b>), Digital Video Recording viewing activity (<b>DVR</b>), Video On Demand viewing activity (<b>VOD</b>), Educational viewing activity (<b>EDU</b>), or Live viewing activity (<b>LIV</b>). The Video Viewing Activity Data can hold all of these formats; for certain formats, some fields will be left blank as determined by the analyst.</p> <p>There are numerous ways to enrich this data. I will teach several enrichment strategies so that an analyst can choose one or more of them to apply to his particular data set.</p> <p>The Video Viewing Activity Data may encompass whatever viewing activity the analyst needs to analyze; this may cover a period of time (an entire day, or primetime viewing), an individual program, DVR viewing, VOD viewing, Educational viewing, Live viewing, etc.</p> <p>Note: The list of fields shown below is a superset of fields meaning that it contains fields that may be used in Linear Tuning Activity, Linear Viewing Activity, Digital Video Recording Playback Activity, Video On Demand Activity, Live viewing, and Educational viewing. Not all of these fields will be used in every one of these viewing types.</p>

**FIG. 5 - Continued**  
**Data Structure**

1000	1002	1004	1006
NAME	TYPE	SIZE	DESCRIPTION
<p><b>Notes:</b> By putting the Program Info <b>1010</b> in this record, this enables the Translation Processors (<b>150 and 154</b>) to translate the value from a detailed value to an aggregated value prior to writing the Program Info to the Prepared Video Viewing Activity Data File <b>160</b> for use by downstream processing.</p>			
<p><b>Notes:</b> By putting the Channel Info <b>1030</b> in this record, this enables the Translation Processors (<b>150 and 154</b>) to translate the value from a detailed value to an aggregated value prior to writing the Channel Info to the Prepared Video Viewing Detail File <b>160</b> for use by downstream processing.</p>			
<p><b>Notes:</b> By Putting the House Info <b>1040</b> in this record, this enables the Translation Processors (<b>150 and 154</b>) to translate the value from a detailed value to an aggregated value prior to writing the House Info to the Prepared Video Viewing Activity Data File <b>160</b> for use by downstream processing.</p>			
<p><b>Notes:</b> By Putting the Device Info <b>1050</b> in this record, this enables the Translation Processors (<b>150 and 154</b>) to translate the value from a detailed value to an aggregated value prior to writing the Device Info to the Prepared Video Viewing Activity Data File <b>160</b> for use by downstream processing.</p>			
<p><b>Notes:</b> By Putting the the Viewer Info <b>1060</b> in this record, this enables the Translation Processors (<b>150 and 154</b>) to translate the value from a detailed value to an aggregated value prior to writing the Viewer Info to the Prepared Video Viewing Activity Data File <b>160</b> for use by downstream processing.</p>			
<p><b>Notes:</b> By Putting the Geographic Information <b>1070</b> in this record, this enables the Translation Processors (<b>150 and 154</b>) to translate the value from a detailed value to an aggregated value prior to writing the Geographic Info to the Prepared Video Viewing Activity Data File <b>160</b> for use by downstream processing.</p>			
<p><b>Notes:</b> By Putting the Demographic Information <b>1080</b> in this record, this enables the Translation Processors (<b>150 and 154</b>) to translate the value from a detailed value to an aggregated value prior to writing the Demographic Info to the Prepared Video Viewing Activity Data File <b>160</b> for use by downstream processing.</p>			

**FIG. 5 – Continued**  
**Data Structure**

**Note:** Those skilled in the art will readily be able to identify the entities in the Cable Television Laboratories, Inc. Media Measurement Data Model which can be used to populate the fields listed below.

1010~	PROGRAM__INFO	CHAR	40	<p>The Program Information field records information about the program. As a non-limiting example, it may contain any of the following:</p> <ul style="list-style-type: none"><li>a) Program identifier</li><li>b) Program type (news, sports, sitcom, etc.)</li><li>c) Program genre (romance, mystery, etc.)</li><li>d) Program name or partial name</li><li>e) Program rating (G, PG, PG17, etc.)</li><li>f) Program duration</li><li>g) Program series flag (Y or N)</li><li>h) Program season code</li><li>i) Program episode number</li><li>j) An online learning course video identifier</li><li>k) Streaming video identifier</li></ul>
1020~	PROGRAM__AIRING__INFO	CHAR	40	<p>The Program Airing Information field records information about an airing of a program. As a non-limiting example, it may contain any of the following:</p> <ul style="list-style-type: none"><li>a) Program airing identifier</li><li>b) Program airing begin date and time</li><li>c) Program airing end date and time</li><li>d) Original airing or replay indicator</li><li>e) Program identifier + Program airing begin date and time</li><li>f) A lesson identifier in an online video learning course</li></ul>



FIG 5 - Data Structure - Continued

1030~	CHANNEL_INFO	CHAR	40	<p>The Channel Information field records information about the channel or network on which the video content aired. As a non-limiting example, it may contain any of the following:</p> <ul style="list-style-type: none"><li>a) Actual Channel call sign (this is the network identifier with any High Definition or Standard Definition information; i.e.: DISCHD and DISC)</li><li>b) Channel call sign (this is the network identifier without any High Definition or Standard Definition information; i.e.: DISC)</li><li>c) Channel source identifier (a numeric channel identifier)</li><li>d) Network identifier</li><li>e) Entertainment Identifier Registry (EIDR) video service identifier (See <a href="http://eidr.org/">http://eidr.org/</a>)</li><li>f) The internet uniform resource locator</li><li>g) The web address or internet protocol address</li></ul>
1040~	HOUSE_INFO	CHAR	40	<p>The House Information field records information about a house that is linked with the viewing activity that is being measured. As a non-limiting example, it may contain any of the following:</p> <ul style="list-style-type: none"><li>a) House identifier</li><li>b) Commercial building identifier</li><li>c) Residential, commercial, multi-family, or hotel indicator</li><li>d) Customer service package accessible by the house</li><li>e) School or classroom identifier in an online learning environment</li></ul>

FIG 5 - Data Structure - Continued

1050~	DEVICE__INFO	CHAR	40	<p>The Device Information field records information about a device which was used to view the content for which viewing activity is being measured. As a non-limiting example, it may contain any of the following:</p> <ul style="list-style-type: none"><li>a) Device identifier (media access control address, serial number, etc.)</li><li>b) Device type (TV, smart phone, Internet Protocol TV, set top box, etc.)</li><li>c) Device characteristic (single tuner, dual tuner, manufacturer, high definition, DVR enabled, etc.)</li></ul>
1060~	VIEWER_INFO	CHAR	40	<p>The Viewer Information field records information about the viewer whose viewing activity is being measured. As a non-limiting example, it may contain any of the following:</p> <ul style="list-style-type: none"><li>a) Viewer identifier</li><li>b) Student identifier in an online learning environment</li><li>c) Viewer type (parent, child, student, etc.)</li><li>d) Grade level in an online learning environment</li><li>e) School name in an online learning environment</li></ul> <p>There may be multiple viewers associated with a house. This can identify a specific viewer in the house. Alternatively, it can identify an attribute applicable to multiple viewers.</p>

FIG 5 - Data Structure - Continued

1070~	<b>GEOGRAPHIC_INFORMATION</b>	CHAR	40	<p>The Geographic Information field records information about the geolocation where the viewing activity that is being measured occurred. As a non-limiting example, it may contain any of the following:</p> <ul style="list-style-type: none"><li>a) A marketing area</li><li>b) A zip code</li><li>c) A cable operator region</li><li>d) A state</li><li>e) A county</li><li>f) A census tract</li><li>g) Cell phone tower identifier</li><li>h) Advertising zone</li><li>i) Satellite television service area</li></ul> <p>See <b>Geographic information</b> in the <b>DEFINITIONS</b> section for a description of the kinds of geographic values that may be linked with a viewing activity.</p>
1080~	<b>DEMOGRAPHIC_INFORMATION</b>	CHAR	40	<p>The Demographic Information field records demographic data about the viewer whose viewing activity is being measured. As a non-limiting example, it may contain any of the following:</p> <ul style="list-style-type: none"><li>a) Age</li><li>b) Income level</li><li>c) Education level</li><li>d) Languages spoken</li><li>e) National origin</li><li>f) Occupation</li><li>g) Home value</li><li>h) Viewing interests</li><li>i) Gender</li><li>j) Marital status</li><li>k) Ages of children in home</li><li>l) Proprietary demographic grouping</li></ul> <p>See <b>Demographic information</b> in the <b>DEFINITIONS</b> section for a description of the kinds of demographic values that may be assigned to a viewer.</p>

FIG 5 - Data Structure - Continued

1090~	VIEWING__TYPE	CHAR	3	The Viewing Type records the the type of viewing activity which is being measured. As a non-limiting example, it may contain any of the following: a) LTA for linear tuning activity b) LVA for sessionized linear viewing activity c) DVR for content being played back from a digital video recording. d) VOD for content being played from a Video On Demand lease e) EDU for material from an online educational course f) LIV for live programming content
1100~	VIEWING__DATE	DATE	8	The Viewing Date records the date of the viewing activity.
1102~	TUNE__IN__DATE__TIME	DATE TIME		Tune in date time records the tune in date and time of the viewing activity.
1104~	TUNE__IN__SECOND__OF__DAY	Numbe r	5	Tune In Second of Day records time of the tune in represented as seconds past midnight.
1106~	TUNE__OUT__DATE__TIME	DATE TIME		Tune out date and time records the tune out date and time of the viewing activity.
1108~	TUNE__OUT__SECON D__OF__DAY	Numbe r	5	Tune out Second of Day records time of the tune out represented as seconds past midnight.

FIG 5 - Data Structure - Continued

1110~	<b>PLAYBACK_MODE</b>	CHAR	2	<p>The Playback Mode records information about the speed and direction of the video content during viewing. Suggested values include:</p> <ul style="list-style-type: none"><li>a) PL for Play</li><li>b) F1 for Fast Forward at 1x speed</li><li>c) F2 for Fast Forward at 2x speed</li><li>d) F3 for Fast Forward at 3x speed</li><li>e) F4 for Fast Forward at 4x speed</li><li>f) R1 for Reverse at 1x speed</li><li>g) R2 for Reverse at 2x speed</li><li>h) R3 for Reverse at 3x speed</li><li>i) R4 for Reverse at 4x speed</li><li>j) PA for Pause</li><li>k) ST for Stop</li><li>l) RE for Resume</li></ul>
1120~	<b>PLAYBACK_BEG_POSITION</b>	Number	5	<p>The Playback Begin Position records the starting point (beginning) of the viewing activity represented by PLAYBACK_MODE based on the actual location in the content from a fixed reference point in the content. The fixed reference point is normally the first second (or frame) of the content (second 0), but it can be any fixed reference point. Having all viewing activity normalized to a fixed reference point enables consistent measurement across many devices operated by many viewers, even when their individual recordings (in the case of DVR) or viewing activity did not begin at the start of the content.</p> <p>This is also referred to as the Beginning Offset Second from the True Beginning.</p>

FIG 5 - Data Structure - Continued

1130~	<b>PLAYBACK_END__ POSITION</b>	Number	5	<p>The Playback End Position records the ending point of the viewing activity represented by PLAYBACK_MODE based on the actual location in the content from a fixed reference point in the content. The fixed reference point is normally the first second (or frame) of the content (second 0), but it can be any fixed reference point. Having all viewing activity normalized to a fixed reference point enables consistent measurement across many devices operated by many viewers, even when their individual recordings (in the case of DVR) or viewing activity did not begin at the start of the content.</p> <p>This is also referred to as the Ending Offset Second from the True Beginning.</p>
1138~	<b>TUNE_DURATION__ SECONDS</b>	Number	5	Tune Duration Seconds records the duration of the tuning event in seconds.
1140~	<b>ACTIVITY_DURATION__ SECONDS</b>	Number	5	The Activity Duration field records the duration of the activity represented by the Playback Mode.
1150~	<b>PRESENTATION__I NFO</b>	CHAR	4	<p>The Presentation Information field records information about the presentation of the video content for which viewing activity is being measured. As a non-limiting example, it may contain any of the following:</p> <ul style="list-style-type: none"> <li>a) Volume level</li> <li>b) Muting</li> <li>c) Closed captioning</li> <li>d) Picture in picture active indicator</li> <li>e) Percent of content visible</li> <li>f) Percent of screen used by content</li> </ul>

FIG 5 - Data Structure - Continued

1160~	VIDEO_SERVER_I NFO	CHAR	40	The Video Server Information field records information about the video server that served the video stream for which viewing activity is being measured. As a non-limiting example, it may contain any of the following: a) A computer system identifier b) A headend identifier c) A hub identifier d) A satellite identifier e) A web server f) A content delivery system identifier
1170~	SYSTEM_HEALTH_ INFO	CHAR	40	The System Health Information field records information about the health of the video server that served the video stream for which viewing activity is being measured. As a non-limiting example, it may contain any of the following: a) A system status indicator b) An error code

**FIG. 6**  
**Summary Information**

Specification Name:	<b>Channel Information Translation Table 140 or 142</b>
Record description:	<p>The <b>Channel Information Translation Table</b> contains information needed to translate from a source channel to a corresponding target channel. The <b>Channel Information Translation Table</b> may be presented to the Channel Translation Processor <b>150</b> as either a flat file <b>140</b> which can be loaded to an array in the memory of the computer, or as a data table which can be joined to <b>142</b>.</p> <p>The <b>Channel Translation Processor 150</b> can be written to accept either form of input.</p>

**Data Structure**

	1800	1802	1804	1806
	<b>NAME</b>	<b>TYPE</b>	<b>SIZE</b>	<b>DESCRIPTION</b>
1810~	<b>SOURCE_CHANNEL__INFORMATION__IN</b>	CHAR	40	The Source Channel Information field contains a value which matches the value which is found in the Channel Information field <b>1030</b> in the Video Viewing Activity Data File <b>130</b> . This is the “from” value.
1820~	<b>TARGET_CHANNEL__INFORMATION__IN</b>	CHAR	40	The Target Channel Information field contains the value which is to be used in the aggregation process. The Channel Information field <b>1030</b> in the Video Viewing Activity Data File <b>130</b> can be replaced with this Target Channel Information field or a new field can be appended to the Video Viewing Activity Data File <b>130</b> record containing the Target Channel Information field. Either method will allow the viewing activity in the Prepared Video Viewing Activity Data File to be aggregated by the Target Channel Information field. This is the “to” value. In <b>FIG 3 Step 2</b> I describe appending the translated value to the output file as Channel_Info_Target <b>1400</b> .



**FIG. 6 - Sample Data****Example 1:**

Translating Standard Definition and High Definition Call Signs to a Common Call Sign.

<b>1810</b>	<b>1820</b>
<b>SOURCE_CHANNEL_INFORMATION_IN</b>	<b>TARGET_CHANNEL_INFORMATION_IN</b>
HIST	HIST
HISTHD	HIST
FOX	FOX
FOXHD	FOX

**Example 2:**

Translating selected Standard Definition and High Definition Call Signs to a Common Call Sign AND everything else to OTHER.

<b>1810</b>	<b>1820</b>
<b>SOURCE_CHANNEL_INFORMATION_IN</b>	<b>TARGET_CHANNEL_INFORMATION_IN</b>
HIST	HIST
HISTHD	HIST
DISC	OTHER
ESPN	OTHER
FOX	OTHER

FIG. 7

Summary Information

Specification Name:	<b>Channel Mapping Definition Array Data Structure</b>
Record description:	<p>The <b>Channel Mapping Definition Array Data Structure</b> is the data structure used to record the Channel Mapping Definitions in the memory of the computer. This Data Structure is loaded from the <b>Channel Information Translation Table 140, FIG 6.</b></p> <p><b>Usage:</b> During the data translation process, as part of enriching the Video Viewing data, the Data Translation Processor <b>150</b> looks for a match on the Channel Information in the Video Viewing Activity Record with the value in <b>CHAN-MAP-CHAN-INFO-SOURCE 1910</b>. Then the system enriches the Video Viewing Record with the <b>CHAN-MAP-CHAN-INFO-TARGET 1920</b>.</p>

Data Structure

	1900		1902		1904		1906
	NAME		TYPE		SIZE		DESCRIPTION
	The following fields occur 600 times to allow for 600 Channel Mappings to be processed. This can be changed as needed.						
	<b>Note:</b> Mapping multiple source channels to a single target channel enables the analysis of viewing activity across different channels that carry the same content while maintaining the detail tuning events in the source video viewing activity file for use in specific channel viewing measurements such as bandwidth consumption.						
	<b>Note:</b> The Channel Information field is not limited to channel call signs.						
1910~	CHAN-MAP-CHAN-INFO-SOURCE		CHAR		40		The Source Channel Map Channel Information as represented in the file. Populated from FIG 6 1810.
1920~	CHAN-MAP-CHAN-INFO-TARGET		CHAR		40		The Channel Map Channel Information that should be used for aggregation purposes; the target channel information. Populated from FIG 6 1820.

**FIG. 8**  
**Summary Information**

Specification Name:	<b>Data Translation Table 146 or 148</b>
Record description:	<p>The <b>Data Translation Table</b> contains information needed to translate from a source value to a corresponding target value. The <b>Data Translation Table</b> may be presented to the Data Translation Processor <b>154</b> as either a flat file which can be loaded to an array in the memory of the computer, or as a data table which can be joined to.</p> <p>I have shown this as either a flat file <b>Data Translation Table 146</b> or as a database table <b>Data Translation Table 148</b>.</p> <p>The <b>Data Translation Processor 154</b> can be written to accept either form of input.</p>

**Data Structure**

2000		2002	2004	2006
	<b>NAME</b>	<b>TYPE</b>	<b>SIZE</b>	<b>DESCRIPTION</b>
	<p><b>Notes:</b> By putting <b>TRANSLATION_VALUE_TYPE_IN 2005</b> in this record, this enables the analyst to define any manner of translation values as needed. This provides a great deal of flexibility in mapping from a more detailed set of values to one or more general values which has the effect of making the downstream aggregation process run faster.</p>			
2005~	<b>TRANSLATION_VALUE_TYPE-IN</b>	CHAR	40	<p>The Translation Value Type field contains a value which identifies the kind of data found in the corresponding Source and Target Translation Values. Non-limiting examples include:</p> <ul style="list-style-type: none"> <li>(i) program information,</li> <li>(ii) channel information,</li> <li>(iii) house information,</li> <li>(iv) device information,</li> <li>(v) viewer information,</li> <li>(vi) geographic information,</li> <li>(vii) demographic information.</li> </ul>

FIG 8 - Data Structure – Continued

2010~	SOURCE_TRANSLATION__VALUE-IN	CHAR	40	<p>The Source Translation Value field contains a value which matches the value which is found in the Video Viewing Activity Data File <b>130</b> for the field specified by the field type indicated by the Source Translation Type field. This is the “from” value.</p> <p>The analyst can use any value that will allow him to match on the data values that are present in the Video Viewing Activity Data File <b>130</b>.</p> <p>As a non-limiting examples,</p> <p>when the Source Translation Type represents <b>Program</b> Information, then this Source Translation Value field may contain any value that will be found in field <b>PROGRAM__INFO 1010</b>;</p> <p>when the Source Translation Type represents <b>Channel</b> Information, then this Source Translation Value field may contain any value that will be found in field <b>CHANNEL__INFO 1030</b>;</p> <p>when the Source Translation Type represents <b>House</b> Information, then this Source Translation Value field may contain any value that will be found in field <b>HOUSE__INFO 1040</b>;</p> <p>when the Source Translation Type represents <b>Device</b> Information, then this Source Translation Value field may contain any value that will be found in field <b>DEVICE__INFO 1050</b>;</p> <p>when the Source Translation Type represents <b>Viewer</b> Information, then this Source Translation Value field may contain any value that will be found in field <b>VIEWER__INFO 1060</b>;</p> <p>when the Source Translation Type represents <b>Geographic</b> Information, then this Source Translation Value field may contain any value that will be found in field <b>GEOGRAPHIC__INFO 1070</b>;</p> <p>when the Source Translation Type represents <b>Demographic</b> Information, then this Source Translation Value field may contain any value that will be found in field <b>DEMOGRAPHIC__INFO 1080</b>.</p>
-------	------------------------------	------	----	--

FIG 8 - Data Structure - Continued

2020~	TARGET_TRANSLATION_VALUE-IN	CHAR	40	<p>The Target Translation Value field contains the value which is to be used in the aggregation process. The field in the Video Viewing Activity Data File <b>130</b> can be replaced with this Target Translation Value field or a new field can be appended to the Video Viewing Activity Data File <b>130</b> record containing the Target Translation Value field to create the Prepared Video Viewing Activity File. Either method will allow the viewing activity in the Prepared Video Viewing Activity Data File <b>160</b> to be aggregated by the Target Translation Value field. This is the “to” value.</p> <p>The analyst has a great deal of flexibility in defining the target translation values.</p> <p>As non-limiting example, these values can be used to map like values to a standard value such as would be the case with mapping the Standard Definition and High Definition Channel Call Signs (ex: DISC and DISCHD) to a common Channel Call Sign (ex: DISC).</p> <p>Also as non-limiting example, these values can be used to map detailed values to values that will be used for aggregation. If an analyst needs to compare viewing of sports programming to all other concurrent programming, he may choose to leave any sports content with its original channel identifier and then replace any other content with a channel identifier of ‘OTHER’. This will significantly reduce the load on the MapReduce framework (because the number of unique keys will be much smaller) while providing detailed analysis of the various sports channels.</p> <p>The <b>FIG 8</b> examples below provide numerous examples of data translations.</p>
-------	-----------------------------	------	----	--

**FIG. 8 - Sample Data**

**Example 1:**

Translating Program Information to a Program Type and everything else to OTHER

2005	2010	2020
TRANSLATION__VALUE__TYPE__IN	SOURCE__TRANSLATION__VALUE__IN	TARGET__TRANSLATION__VALUE__IN
PROGRAM	SPORT	SPORT
PROGRAM	DOCU	OTHER
PROGRAM	NATU	OTHER
PROGRAM	NEWS	OTHER

**Example 2:**

Translating Standard Definition and High Definition Channel Call Signs to a Common Channel Call Sign

2005	2010	2020
TRANSLATION__VALUE__TYPE__IN	SOURCE__TRANSLATION__VALUE__IN	TARGET__TRANSLATION__VALUE__IN
CHANNEL	HIST	HIST
CHANNEL	HISTHD	HIST
CHANNEL	FOX	FOX
CHANNEL	FOXHD	FOX

**Example 3:**

Translating selected Standard Definition and High Definition Channel Call Signs to a Common Channel Call Sign AND everything else to OTHER

2005	2010	2020
TRANSLATION__VALUE__TYPE__IN	SOURCE__TRANSLATION__VALUE__IN	TARGET__TRANSLATION__VALUE__IN
CHANNEL	HIST	HIST
CHANNEL	HISTHD	HIST
CHANNEL	DISC	OTHER
CHANNEL	ESPN	OTHER
CHANNEL	FOX	OTHER

**FIG. 8 - Sample Data - Continued****Example 4:**

Translating selected House Information to one value (HOUSE) and everything else to OTHER

2005	2010	2020
TRANSLATION_VALUE_TYPE__ IN	SOURCE_TRANSLATION_VALUE__ IN	TARGET_TRANSLATION_VALUE__ IN
HOUSE	HOUSE	HOUSE
HOUSE	APT	OTHER
HOUSE	OFFICE	OTHER
HOUSE	SCHOOL	OTHER

**Example 5:**

Translating selected Device Information to one value (STB) and everything else to OTHER

2005	2010	2020
TRANSLATION_VALUE_TYPE__ IN	SOURCE_TRANSLATION_VALUE__ IN	TARGET_TRANSLATION_VALUE__ IN
DEVICE	STB	STB
DEVICE	IPTV	OTHER
DEVICE	TABLET	OTHER
DEVICE	PHONE	OTHER

**Example 6:**

Translating selected Viewer Information to one value (PARENT) and everything else to OTHER

2005	2010	2020
TRANSLATION_VALUE_TYPE__ IN	SOURCE_TRANSLATION_VALUE__ IN	TARGET_TRANSLATION_VALUE__ IN
VIEWER	PARENT	PARENT
VIEWER	CHILD	OTHER
VIEWER	STUDENT	OTHER

**FIG. 8 - Sample Data - Continued**

**Example 7:**

Translating Geographic Information from a detailed value to a higher value.

2005	2010	2020
TRANSLATION_VALUE_TYPE_ IN	SOURCE_TRANSLATION_VALUE_ IN	TARGET_TRANSLATION_VALUE_ IN
GEOGRAPHIC	80111	801XX
GEOGRAPHIC	80112	801XX
GEOGRAPHIC	80113	801XX
GEOGRAPHIC	80114	801XX
GEOGRAPHIC	80115	801XX

**Example 8:**

Translating Demographic Information from a detailed value to a higher value

2005	2010	2020
TRANSLATION_VALUE_TYPE_ IN	SOURCE_TRANSLATION_VALUE_ IN	TARGET_TRANSLATION_VALUE_ IN
DEMOGRAPHIC	1-3 YR	CHILD
DEMOGRAPHIC	4-5 YR	CHILD
DEMOGRAPHIC	6-8 YR	CHILD
DEMOGRAPHIC	21-44 YR	ADULT
DEMOGRAPHIC	45-54 YR	ADULT



**FIG. 9**  
**Summary Information**

Specification Name:	<b>Generalized Mapping Definition Array Data Structure</b>
Record description:	<p>The <b>Generalized Mapping Definition Array Data Structure</b> is the data structure used to record the <b>Data Translation Table Mapping Definitions</b> in the memory of the computer. This Data Structure is loaded from the <b>Data Translation Table 146, FIG 8</b>.</p> <p><b>Usage:</b> The Data Translation Processor <b>154</b>, as part of enriching the Video Viewing data, looks for a match against <b>SOURCE_TRANSLATION_VALUE</b> on the field indicated by <b>TRANSLATION_VALUE_TYPE</b>. Then the Processor <b>154</b> builds the Video Viewing Activity Record with the <b>TARGET_TRANSLATION_VALUE</b>. The logic for this process is described in <b>FIG 4</b>.</p>

**Data Structure**

	2100		2102	2104	2106
	NAME	TYPE	SIZE	DESCRIPTION	
	The following fields occur 1000 times to allow for 1000 Data Translation Mappings to be processed. This can be changed as needed.				
	<p><b>Note:</b> Providing multiple translation types with their corresponding source and target translation values allows the analyst to translate multiple values in preparation for aggregation by the MapReduce framework. As a non-limiting example, he may translate both channel information and demographic information to prepare for a particular analytical study. The original source data remains unchanged so that it can be used in any number of studies, some of which may require the original values.</p>				
2105~	TRANSLATION__VALUE__TYPE	CHAR	40	The Translation Value Type. Populated from <b>FIG 8 2005</b> .	
2110~	SOURCE_TRANSLATION__VALUE	CHAR	40	The Source Translation Value. Populated from <b>FIG 8 2010</b> .	
2120~	TARGET_TRANSLATION__VALUE	CHAR	40	The Target Translation Value. Populated from <b>FIG 8 2020</b> .	

FIG. 10 - Sample Data

## Example 1:

Sample Data for LTA showing Geographic Information Translation from a detailed value to a higher value

## Geographic Information - Before

GEO-GRAPHIC INFO	VIDEO SERVER INFO	CHANNEL INFO	VIEW-ING DEVICE	HOUSE	VIEWER	DEMO-GRAPHIC INFO	TUNE-IN DATE TIME	TUNE-IN SECOND	TUNE-OUT DATE TIME	TUNE-OUT SECOND	TUNE DURATION
1070	1160	1030	1050	1040	1060	1080	1102	1104	1106	1108	1138
80111	SERVER-01	ABC	1	H1	U1	40-60k	2013-08-01 18:00:00	64800	2013-08-01 18:00:05	64805	6
80112	SERVER-01	ABC	2	H2	U2	40-60k	2013-08-01 18:00:01	68401	2013-08-01 18:00:07	68407	7
80113	SERVER-01	DISC	3	H3	U3	60-80K	2013-08-01 18:00:02	68402	2013-08-01 18:00:06	68406	5
80114	SERVER-01	DISC	3	H3	U3	60-80K	2013-08-01 18:00:04	68404	2013-08-01 18:00:06	68406	3
80115	SERVER-01	DISC	4	H4	U4	80-110K	2013-08-01 18:00:03	68403	2013-08-01 18:00:09	68409	7
80111	SERVER-01	ESPN	5	H5	U5	80-110K	2013-08-01 18:00:04	68404	2013-08-01 18:00:11	68411	8

## Geographic Information - After

GEO-GRAPHIC INFO	VIDEO SERVER INFO	CHANNEL INFO	VIEW-ING DEVICE	HOUSE	VIEWER	DEMO-GRAPHIC INFO	TUNE-IN DATE TIME	TUNE-IN SECOND	TUNE-OUT DATE TIME	TUNE-OUT SECOND	TUNE DURATION
1270	1360	1230	1250	1240	1260	1280	1302	1304	1306	1308	1338
801XX	SERVER-01	ABC	1	H1	U1	40-60k	2013-08-01 18:00:00	64800	2013-08-01 18:00:05	64805	6
801XX	SERVER-01	ABC	2	H2	U2	40-60k	2013-08-01 18:00:01	68401	2013-08-01 18:00:07	68407	7
801XX	SERVER-01	DISC	3	H3	U3	60-80K	2013-08-01 18:00:02	68402	2013-08-01 18:00:06	68406	5
801XX	SERVER-01	DISC	3	H3	U3	60-80K	2013-08-01 18:00:04	68404	2013-08-01 18:00:06	68406	3
801XX	SERVER-01	DISC	4	H4	U4	80-110K	2013-08-01 18:00:03	68403	2013-08-01 18:00:09	68409	7
801XX	SERVER-01	ESPN	5	H5	U5	80-110K	2013-08-01 18:00:04	68404	2013-08-01 18:00:11	68411	8

FIG. 10 - Sample Data – Continued

## Example 2:

## Sample Data for LTA showing Channel Information Translation

Translating Channel Identifier to a Common Channel Identifier from normal and HD versions of the channel

## Channel Information - Before

GEO-GRAPHIC INFO	VIDEO SERVER INFO	CHANNEL INFO	VIEWING DEVICE	HOUSE	VIEWER	DEMO-GRAPHIC INFO	TUNE-IN DATE TIME	TUNE-IN SECOND	TUNE-OUT DATE TIME	TUNE-OUT SECOND	TUNE DURATION
1070	1160	1030	1050	1040	1060	1080	1102	1104	1106	1108	1138
80111	SERVER-01	HIST	1	H1	U1	40-60k	2013-08-01 18:00:00	64800	2013-08-01 18:00:05	64805	6
80112	SERVER-01	HISTHD	2	H2	U2	40-60k	2013-08-01 18:00:01	68401	2013-08-01 18:00:07	68407	7
80113	SERVER-01	FOX	3	H3	U3	60-80K	2013-08-01 18:00:02	68402	2013-08-01 18:00:06	68406	5
80114	SERVER-01	FOXHD	3	H3	U3	60-80K	2013-08-01 18:00:04	68404	2013-08-01 18:00:06	68406	3

## Channel Information - After

GEO-GRAPHIC INFO	VIDEO SERVER INFO	CHANNEL INFO	VIEWING DEVICE	HOUSE	VIEWER	DEMO-GRAPHIC INFO	TUNE-IN DATE TIME	TUNE-IN SECOND	TUNE-OUT DATE TIME	TUNE-OUT SECOND	TUNE DURATION
1270	1360	1230	1250	1240	1260	1280	1302	1304	1306	1308	1338
80111	SERVER-01	HIST	1	H1	U1	40-60k	2013-08-01 18:00:00	64800	2013-08-01 18:00:05	64805	6
80112	SERVER-01	HIST	2	H2	U2	40-60k	2013-08-01 18:00:01	68401	2013-08-01 18:00:07	68407	7
80113	SERVER-01	FOX	3	H3	U3	60-80K	2013-08-01 18:00:02	68402	2013-08-01 18:00:06	68406	5
80114	SERVER-01	FOX	3	H3	U3	60-80K	2013-08-01 18:00:04	68404	2013-08-01 18:00:06	68406	3

FIG. 10 - Sample Data – Continued

## Example 3:

Sample Data for LTA showing Channel Information Translation

Channel Example – Translating Channel Identifier to a Common Channel Identifier and All OTHER

## Channel Information - Before

GEO- GRAPHIC INFO	VIDEO SERVER INFO	CHAN- NEL INFO	VIEW- ING DEV ICE	HOUSE	VIEW ER	DEMO- GRAPHIC INFO	TUNE-IN DATE TIME	TUNE-IN SECOND	TUNE-OUT DATE TIME	TUNE- OUT SECOND	TUNE DURAT ION
1070	1160	1030	1050	1040	1060	1080	1102	1104	1106	1108	1138
80111	SERVER-01	HIST	1	H1	U1	40-60k	2013-08-01 18:00:00	64800	2013-08-01 18:00:05	64805	6
80112	SERVER-01	HISTHD	2	H2	U2	40-60k	2013-08-01 18:00:01	68401	2013-08-01 18:00:07	68407	7
80113	SERVER-01	DJSC	3	H3	U3	60-80K	2013-08-01 18:00:02	68402	2013-08-01 18:00:06	68406	5
80114	SERVER-01	ESPN	3	H3	U3	60-80K	2013-08-01 18:00:04	68404	2013-08-01 18:00:06	68406	3
801XX	SERVER-01	FOX	4	H4	U4	80-110K	2013-08-01 18:00:03	68403	2013-08-01 18:00:09	68409	7

## Channel Information - After

GEO- GRAPHIC INFO	VIDEO SERVER INFO	CHAN- NEL INFO	VIEW- ING DEV ICE	HOUSE	VIEW ER	DEMO- GRAPHIC INFO	TUNE-IN DATE TIME	TUNE-IN SECOND	TUNE-OUT DATE TIME	TUNE- OUT SECOND	TUNE DURAT ION
1270	1360	1230	1250	1240	1260	1280	1302	1304	1306	1308	1338
80111	SERVER-01	HIST	1	H1	U1	40-60k	2013-08-01 18:00:00	64800	2013-08-01 18:00:05	64805	6
80112	SERVER-01	HIST	2	H2	U2	40-60k	2013-08-01 18:00:01	68401	2013-08-01 18:00:07	68407	7
80113	SERVER-01	OTHER	3	H3	U3	60-80K	2013-08-01 18:00:02	68402	2013-08-01 18:00:06	68406	5
80114	SERVER-01	OTHER	3	H3	U3	60-80K	2013-08-01 18:00:04	68404	2013-08-01 18:00:06	68406	3
801XX	SERVER-01	OTHER	4	H4	U4	80-110K	2013-08-01 18:00:03	68403	2013-08-01 18:00:09	68409	7

FIG. 10 - Sample Data – Continued

## Example 4:

Sample Data for LTA showing Device Information Translation

Device Example – Translating Device Identifier to Common Device Identifier and All OTHER

## Device Information - Before

GEO- GRAPHIC INFO	VIDEO SERVER INFO	CHAN- NEL INFO	VIEW- ING DEV ICE	HOUSE	VIEW ER	DEMO- GRAPHIC INFO	TUNE-IN DATE TIME	TUNE-IN SECOND	TUNE-OUT DATE TIME	TUNE- OUT SECOND	TUNE DURAT ION
1070	1160	1030	1050	1040	1060	1080	1102	1104	1106	1108	1138
80111	SERVER-01	HIST	STB	HOUSE	U1	40-60k	2013-08-01 18:00:00	64800	2013-08-01 18:00:05	64805	6
80112	SERVER-01	HISTHD	IPTV	APT	U2	40-60k	2013-08-01 18:00:01	68401	2013-08-01 18:00:07	68407	7
80113	SERVER-01	DISC	TABLET	OFFICE	U3	60-80K	2013-08-01 18:00:02	68402	2013-08-01 18:00:06	68406	5
80114	SERVER-01	ESPN	PHONE	SCHOOL	U3	60-80K	2013-08-01 18:00:04	68404	2013-08-01 18:00:06	68406	3
801XX	SERVER-01	FOX	STB	HOUSE	U4	80-110K	2013-08-01 18:00:03	68403	2013-08-01 18:00:09	68409	7

## Device Information - After

GEO- GRAPHIC INFO	VIDEO SERVER INFO	CHAN- NEL INFO	VIEW- ING DEV ICE	HOUSE	VIEW ER	DEMO- GRAPHIC INFO	TUNE-IN DATE TIME	TUNE-IN SECOND	TUNE-OUT DATE TIME	TUNE- OUT SECOND	TUNE DURAT ION
1270	1360	1230	1250	1240	1260	1280	1302	1304	1306	1308	1338
80111	SERVER-01	HIST	STB	HOUSE	U1	40-60k	2013-08-01 18:00:00	64800	2013-08-01 18:00:05	64805	6
80112	SERVER-01	HISTHD	OTHER	APT	U2	40-60k	2013-08-01 18:00:01	68401	2013-08-01 18:00:07	68407	7
80113	SERVER-01	DISC	OTHER	OFFICE	U3	60-80K	2013-08-01 18:00:02	68402	2013-08-01 18:00:06	68406	5
80114	SERVER-01	ESPN	OTHER	SCHOOL	U3	60-80K	2013-08-01 18:00:04	68404	2013-08-01 18:00:06	68406	3
801XX	SERVER-01	FOX	STB	HOUSE	U4	80-110K	2013-08-01 18:00:03	68403	2013-08-01 18:00:09	68409	7

FIG. 10 - Sample Data – Continued

Example 5:

Sample Data for LTA showing House Information Translation

House Example – Translating House Identifier to Common House Identifier and All OTHER

House Information - Before

GEO- GRAPHIC INFO	VIDEO SERVER INFO	CHAN- NEL INFO	VIEW- ING DEV ICE	HOUSE	VIEW ER	DEMO- GRAPHIC INFO	TUNE-IN DATE TIME	TUNE-IN SECOND	TUNE-OUT DATE TIME	TUNE- OUT SECOND	TUNE DURAT ION
1070	1160	1030	1050	1040	1060	1080	1102	1104	1106	1108	1138
80111	SERVER- 01	HIST	1	HOUSE	U1	40-60k	2013-08-01 18:00:00	64800	2013-08-01 18:00:05	64805	6
80112	SERVER- 01	HISTHD	2	APT	U2	40-60k	2013-08-01 18:00:01	68401	2013-08-01 18:00:07	68407	7
80113	SERVER- 01	DISC	3	OFFICE	U3	60-80K	2013-08-01 18:00:02	68402	2013-08-01 18:00:06	68406	5
80114	SERVER- 01	ESPN	3	SCHOOL	U3	60-80K	2013-08-01 18:00:04	68404	2013-08-01 18:00:06	68406	3
801XX	SERVER- 01	FOX	4	HOUSE	U4	80-110K	2013-08-01 18:00:03	68403	2013-08-01 18:00:09	68409	7

House Information - After

GEO- GRAPHIC INFO	VIDEO SERVER INFO	CHAN- NEL INFO	VIEW- ING DEV ICE	HOUSE	VIEW ER	DEMO- GRAPHIC INFO	TUNE-IN DATE TIME	TUNE-IN SECOND	TUNE-OUT DATE TIME	TUNE- OUT SECOND	TUNE DURAT ION
1270	1360	1230	1250	1240	1260	1280	1302	1304	1306	1308	1338
80111	SERVER- 01	HIST	1	HOUSE	U1	40-60k	2013-08-01 18:00:00	64800	2013-08-01 18:00:05	64805	6
80112	SERVER- 01	HISTHD	2	OTHER	U2	40-60k	2013-08-01 18:00:01	68401	2013-08-01 18:00:07	68407	7
80113	SERVER- 01	DISC	3	OTHER	U3	60-80K	2013-08-01 18:00:02	68402	2013-08-01 18:00:06	68406	5
80114	SERVER- 01	ESPN	3	OTHER	U3	60-80K	2013-08-01 18:00:04	68404	2013-08-01 18:00:06	68406	3
801XX	SERVER- 01	FOX	4	HOUSE	U4	80-110K	2013-08-01 18:00:03	68403	2013-08-01 18:00:09	68409	7

FIG. 10 - Sample Data -- Continued

Example 6:

Sample Data for LTA showing Viewer Information Translation

Viewer Example -- Translating Viewer Identifier to Common Viewer Identifier and All OTHER

Viewer Information - Before

GEO- GRAPHIC INFO	VIDEO SERVER INFO	CHAN- NEL INFO	VIEW- ING DEV ICE	HOUSE	VIEW ER	DEMO- GRAPHIC INFO	TUNE-IN DATE TIME	TUNE-IN SECOND	TUNE-OUT DATE TIME	TUNE- OUT SECOND	TUNE DURAT ION
1070	1160	1030	1050	1040	1060	1080	1102	1104	1106	1108	1138
80111	SERVER- 01	HIST	1	HOUSE	PARENT	40-60k	2013-08-01 18:00:00	64800	2013-08-01 18:00:05	64805	6
80112	SERVER- 01	HISTHD	2	APT	CHILD	40-60k	2013-08-01 18:00:01	68401	2013-08-01 18:00:07	68407	7
80113	SERVER- 01	DISC	3	OFFICE	STUDENT	60-80K	2013-08-01 18:00:02	68402	2013-08-01 18:00:06	68406	5
80114	SERVER- 01	ESPN	3	SCHOOL	PARENT	60-80K	2013-08-01 18:00:04	68404	2013-08-01 18:00:06	68406	3
801XX	SERVER- 01	FOX	4	HOUSE	CHILD	80-110K	2013-08-01 18:00:03	68403	2013-08-01 18:00:09	68409	7

Viewer Information - After

GEO- GRAPHIC INFO	VIDEO SERVER INFO	CHAN- NEL INFO	VIEW- ING DEV ICE	HOUSE	VIEW ER	DEMO- GRAPHIC INFO	TUNE-IN DATE TIME	TUNE-IN SECOND	TUNE-OUT DATE TIME	TUNE- OUT SECOND	TUNE DURAT ION
1270	1360	1230	1250	1240	1260	1280	1302	1304	1306	1308	1338
80111	SERVER- 01	HIST	1	HOUSE	PARENT	40-60k	2013-08-01 18:00:00	64800	2013-08-01 18:00:05	64805	6
80112	SERVER- 01	HISTHD	2	APT	OTHER	40-60k	2013-08-01 18:00:01	68401	2013-08-01 18:00:07	68407	7
80113	SERVER- 01	DISC	3	OFFICE	OTHER	60-80K	2013-08-01 18:00:02	68402	2013-08-01 18:00:06	68406	5
80114	SERVER- 01	ESPN	3	SCHOOL	PARENT	60-80K	2013-08-01 18:00:04	68404	2013-08-01 18:00:06	68406	3
801XX	SERVER- 01	FOX	4	HOUSE	OTHER	80-110K	2013-08-01 18:00:03	68403	2013-08-01 18:00:09	68409	7

FIG. 10 - Sample Data – Continued

Example 7:

Sample Data for LTA showing Demographic Information Translation

Demographic Example – Translating Demographic Information from a detailed value to a higher value  
OTHER

Demographic Information - Before

GEO- GRAPHIC INFO	VIDEO SERVER INFO	CHAN- NEL INFO	VIEW- ING DEV ICE	HOUSE	VIEW ER	DEMO- GRAPHIC INFO	TUNE-IN DATE TIME	TUNE-IN SECOND	TUNE-OUT DATE TIME	TUNE- OUT SECOND	TUNE DURAT ION
1070	1160	1030	1050	1040	1060	1080	1102	1104	1106	1108	1138
80111	SERVER-01	HIST	1	HOUSE	U1	1-3 YR	2013-08-01 18:00:00	64800	2013-08-01 18:00:05	64805	6
80112	SERVER-01	HISTHD	2	APT	U2	4-5 YR	2013-08-01 18:00:01	68401	2013-08-01 18:00:07	68407	7
80113	SERVER-01	DISC	3	OFFICE	U3	6-8 YR	2013-08-01 18:00:02	68402	2013-08-01 18:00:06	68406	5
80114	SERVER-01	ESPN	3	SCHOOL	U4	4-5 YR	2013-08-01 18:00:04	68404	2013-08-01 18:00:06	68406	3
801XX	SERVER-01	FOX	4	HOUSE	U5	21-44 YR	2013-08-01 18:00:03	68403	2013-08-01 18:00:09	68409	7
801XX	SERVER-01	FOX	4	HOUSE	U6	45-54 YR	2013-08-01 18:00:03	68403	2013-08-01 18:00:09	68409	7

Demographic Information - After

GEO- GRAPHIC INFO	VIDEO SERVER INFO	CHAN- NEL INFO	VIEW- ING DEV ICE	HOUSE	VIEW ER	DEMO- GRAPHIC INFO	TUNE-IN DATE TIME	TUNE-IN SECOND	TUNE-OUT DATE TIME	TUNE- OUT SECOND	TUNE DURAT ION
1270	1360	1230	1250	1240	1260	1280	1302	1304	1306	1308	1338
80111	SERVER-01	HIST	1	HOUSE	U1	CHILD	2013-08-01 18:00:00	64800	2013-08-01 18:00:05	64805	6
80112	SERVER-01	HISTHD	2	APT	U2	CHILD	2013-08-01 18:00:01	68401	2013-08-01 18:00:07	68407	7
80113	SERVER-01	DISC	3	OFFICE	U3	CHILD	2013-08-01 18:00:02	68402	2013-08-01 18:00:06	68406	5
80114	SERVER-01	ESPN	3	SCHOOL	U4	CHILD	2013-08-01 18:00:04	68404	2013-08-01 18:00:06	68406	3
801XX	SERVER-01	FOX	4	HOUSE	U5	ADULT	2013-08-01 18:00:03	68403	2013-08-01 18:00:09	68409	7
801XX	SERVER-01	FOX	4	HOUSE	U6	ADULT	2013-08-01 18:00:03	68403	2013-08-01 18:00:09	68409	7



FIG. 11 - Sample Data

Example 1:

Sample Data for LVA, DVR, VOD, EDU, LIV showing Program Information Translation

Program Information – Before

PRO-GRAM	PROGRAM AIRING	CHAN-NEL INFO	HOUSE	VIEWING DEVICE	VIEWER	GEO-GRAPHIC INFO	DEMO-GRAPHIC INFO	VIEWING TYPE	VIEWING DATE	PLAY-BACK MODE	PLAY-BACK BEGIN POSITION	PLAY-BACK END POSITION	ACTIVITY DURATION SECONDS	ASSESSMENT-ATION INFO	VIDEO SERVER INFO	SYSTEM HEALTH INFO
1010	1020	1030	1040	1050	1060	1070	1080	1090	1100	1110	1120	1130	1140	1150	1160	1170
SPORT	1001-A	ESPN	H21	D21	V21	DENV	40-60k	LIN	2013-08-31	PL	3	7	5	NORM	SRV-01	A+
DOCU	1002-A	HIST	H22	D22	V22	DENV	40-60k	LIN	2013-08-31	PL	2	6	5	NORM	SRV-01	A+
NATU	1003-A	HIST	H23	D23	V23	DENV	40-60k	LIN	2013-08-31	PL	1	7	7	NORM	SRV-01	A+
NEWS	1004-A	ABC	H24	D24	V24	DENV	40-60k	LIN	2013-08-31	PL	3	6	4	NORM	SRV-01	A+
SPORT	1005-A	CBS	H25	D25	V25	DENV	40-60k	LIN	2013-08-31	PL	4	12	9	NORM	SRV-01	A+

Program Information – After

PRO-GRAM	PROGRAM AIRING	CHAN-NEL INFO	HOUSE	VIEWING DEVICE	VIEWER	GEO-GRAPHIC INFO	DEMO-GRAPHIC INFO	VIEWING TYPE	VIEWING DATE	PLAY-BACK MODE	PLAY-BACK BEGIN POSITION	PLAY-BACK END POSITION	ACTIVITY DURATION SECONDS	ASSESSMENT-ATION INFO	VIDEO SERVER INFO	SYSTEM HEALTH INFO
1210	1220	1230	1240	1250	1260	1270	1280	1290	1300	1310	1320	1330	1340	1350	1360	1370
SPORT	1001-A	ESPN	H21	D21	V21	DENV	40-60k	LIN	2013-08-31	PL	3	7	5	NORM	SRV-01	A+
OTHER	1002-A	HIST	H22	D22	V22	DENV	40-60k	LIN	2013-08-31	PL	2	6	5	NORM	SRV-01	A+
OTHER	1003-A	HIST	H23	D23	V23	DENV	40-60k	LIN	2013-08-31	PL	1	7	7	NORM	SRV-01	A+
OTHER	1004-A	HIST	H24	D24	V24	DENV	40-60k	LIN	2013-08-31	PL	3	6	4	NORM	SRV-01	A+
SPORT	1005-A	CBS	H25	D25	V25	DENV	40-60k	LIN	2013-08-31	PL	4	12	9	NORM	SRV-01	A+

FIG. 11 - Sample Data – Continued

Example 2:

Sample Data for Linear Viewing Activity, DVR, VOD, EDU, LIV

Channel Example – Translating Channel Identifier to a Common Channel Identifier from normal and HD versions of the channel

Channel Information – Before

PRO-GRAM	PROGRAM AIRING	CHAN-NEL INFO	HOUSE	VIEWING DEVICE	VIEWER	GEO-GRAPHIC INFO	DEMO-GRAPHIC INFO	VIEWING TYPE	VIEWING DATE	PLAY-BACK MODE	PLAY-BACK BEGIN POSITION	PLAY-BACK END POSITION	ACTIVITY DURATION RECORD	PRESENCE-ACTION INFO	VIDEO SERVER INFO	SYSTEM HEALTH INFO
1010	1020	1030	1040	1050	1060	1070	1080	1090	1100	1110	1120	1130	1140	1150	1160	1170
1001	1001-A	HIST	H20	D20	V21	DENV	40-60k	LIN	2013-08-31	PL	3	7	5	NORM	SRV-01	A+
1002	1002-A	HISTHD	H21	D21	V21	DENV	40-60k	LIN	2013-08-31	PL	2	6	5	NORM	SRV-01	A+
1003	1003-A	FOX	H23	D23	V23	DENV	40-60k	LIN	2013-08-31	PL	3	7	5	NORM	SRV-01	A+
1004	1004-A	FOXHD	H24	D24	V24	DENV	40-60k	LIN	2013-08-31	PL	2	6	5	NORM	SRV-01	A+

Channel Information – After

PRO-GRAM	PROGRAM AIRING	CHAN-NEL INFO	HOUSE	VIEWING DEVICE	VIEWER	GEO-GRAPHIC INFO	DEMO-GRAPHIC INFO	VIEWING TYPE	VIEWING DATE	PLAY-BACK MODE	PLAY-BACK BEGIN POSITION	PLAY-BACK END POSITION	ACTIVITY DURATION RECORD	PRESENCE-ACTION INFO	VIDEO SERVER INFO	SYSTEM HEALTH INFO
1210	1220	1230	1240	1250	1260	1270	1280	1290	1300	1310	1320	1330	1340	1350	1360	1370
1001	1001-A	HIST	H20	D20	V21	DENV	40-60k	LIN	2013-08-31	PL	3	7	5	NORM	SRV-01	A+
1002	1002-A	HIST	H21	D21	V21	DENV	40-60k	LIN	2013-08-31	PL	2	6	5	NORM	SRV-01	A+
1003	1003-A	FOX	H23	D23	V23	DENV	40-60k	LIN	2013-08-31	PL	3	7	5	NORM	SRV-01	A+
1004	1004-A	FOX	H24	D24	V24	DENV	40-60k	LIN	2013-08-31	PL	2	6	5	NORM	SRV-01	A+

FIG. 11 - Sample Data – Continued

## Example 3:

Sample Data for Linear Viewing Activity, DVR, VOD, EDU, LIV

Channel Example – Translating Channel Identifier to a Common Channel Identifier and All OTHER

## Channel Information – Before

PRO-GRAM	PROGRAM AIRING	CHAN-NEL INFO	HOUSE	VIEWING SERVICE	VIEWER	GEO-GRAPHIC INFO	DEMO-GRAPHIC INFO	VIEWING TYPE	VIEWING DATE	PLAY-BACK MODE	PLAY-BACK BEGIN POSITION	PLAY-BACK END POSITION	ACTIVITY INFORMATION RECORD	PERMANENT INFO	VIDEO SERVER INFO	SYSTEM RECORD INFO
1010	1020	1030	1040	1050	1060	1070	1080	1090	1100	1110	1120	1130	1140	1150	1160	1170
PGM1	PGM1AIR1	HIST	H21	D21	V21	DENV	40-60k	DVR	2013-08-31	PL	3	7	5	NORM	SRV-01	A+
PGM1	PGM1AIR1	HISTHD	H22	D22	V22	DENV	40-60k	DVR	2013-08-31	PL	2	6	5	NORM	SRV-01	A+
PGM1	PGM1AIR2	DISC	H23	D23	V23	DENV	80-110K	DVR	2013-08-31	R1	7	11	5	NORM	SRV-01	A+
PGM1	PGM1AIR2	ESPN	H24	D24	V24	DENV	80-110K	DVR	2013-08-31	R2	3	6	4	NORM	SRV-01	A+
PGM1	PGM1AIR2	FOX	H25	D25	V25	DENV	80-110K	DVR	2013-08-31	PL	4	15	12	NORM	SRV-01	A+

## Channel Information – After

PRO-GRAM	PROGRAM AIRING	CHAN-NEL INFO	HOUSE	VIEWING SERVICE	VIEWER	GEO-GRAPHIC INFO	DEMO-GRAPHIC INFO	VIEWING TYPE	VIEWING DATE	PLAY-BACK MODE	PLAY-BACK BEGIN POSITION	PLAY-BACK END POSITION	ACTIVITY INFORMATION RECORD	PERMANENT INFO	VIDEO SERVER INFO	SYSTEM RECORD INFO
1210	1220	1230	1240	1250	1260	1270	1280	1290	1300	1310	1320	1330	1340	1350	1360	1370
PGM1	PGM1AIR1	HIST	H20	D20	V21	DENV	40-60k	DVR	2013-08-31	PL	3	7	5	NORM	SRV-01	A+
PGM1	PGM1AIR1	HIST	H21	D21	V21	DENV	40-60k	DVR	2013-08-31	PL	2	6	5	NORM	SRV-01	A+
PGM1	PGM1AIR2	OTHER	H23	D23	V23	DENV	80-110K	DVR	2013-08-31	R1	7	11	5	NORM	SRV-01	A+
PGM1	PGM1AIR2	OTHER	H24	D24	V24	DENV	80-110K	DVR	2013-08-31	R2	3	6	4	NORM	SRV-01	A+
PGM1	PGM1AIR2	OTHER	H25	D25	V25	DENV	80-110K	DVR	2013-08-31	PL	4	15	12	NORM	SRV-01	A+

FIG. 11 - Sample Data – Continued

Example 4:

Sample Data for Linear Viewing Activity, DVR, VOD, EDU, LIV

House Example – Translating House Identifier to Common House Identifier and All other

House Information – Before

PRO-GRAM	PROGRAM AIRING	CHAN-NEL INFO	HOUSE	VIEWING DEVICE	VIEWER	GEO-GRAPHIC INFO	DEMO-GRAPHIC INFO	VIEWING TYPE	VIEWING DATE	PLAY-BACK MODE	PLAY-BACK BEGIN POSITION	PLAY-BACK END POSITION	ACTIVITY DURATION SECONDS	PRESENTATION INFO	VIDEO SERVER INFO	SYSTEM HEALTH INFO
1010	1020	1030	1040	1050	1060	1070	1080	1090	1100	1110	1120	1130	1140	1150	1160	1170
1001	1001-A	HIST	HOUSE	D20	V21	DENV	40-60k	VOD	2013-08-31	PL	3	7	5	NORM	SRV-01	A+
1002	1002-A	HISTHD	APT	D21	V21	DENV	40-60k	VOD	2013-08-31	PL	2	6	5	NORM	SRV-01	A+
1003	1003-A	FOX	OFFICE	D23	V23	DENV	40-60k	VOD	2013-08-31	PL	3	7	5	NORM	SRV-01	A+
1004	1004-A	FOXHD	SCHOOL	D24	V24	DENV	40-60k	VOD	2013-08-31	PL	2	6	5	NORM	SRV-01	A+
1005	1005-A	DISC	HOUSE	D23	V23	DENV	80-110K	VOD	2013-09-01	PL	4	15	12	NORM	SRV-01	A+

House Information – After

PRO-GRAM	PROGRAM AIRING	CHAN-NEL INFO	HOUSE	VIEWING DEVICE	VIEWER	GEO-GRAPHIC INFO	DEMO-GRAPHIC INFO	VIEWING TYPE	VIEWING DATE	PLAY-BACK MODE	PLAY-BACK BEGIN POSITION	PLAY-BACK END POSITION	ACTIVITY DURATION SECONDS	PRESENTATION INFO	VIDEO SERVER INFO	SYSTEM HEALTH INFO
1210	1220	1230	1240	1250	1260	1270	1280	1290	1300	1310	1320	1330	1340	1350	1360	1370
1001	1001-A	HIST	HOUSE	D20	V21	DENV	40-60k	VOD	2013-08-31	PL	3	7	5	NORM	SRV-01	A+
1002	1002-A	HISTHD	OTHER	D21	V21	DENV	40-60k	VOD	2013-08-31	PL	2	6	5	NORM	SRV-01	A+
1003	1003-A	FOX	OTHER	D23	V23	DENV	40-60k	VOD	2013-08-31	PL	3	7	5	NORM	SRV-01	A+
1004	1004-A	FOXHD	OTHER	D24	V24	DENV	40-60k	VOD	2013-08-31	PL	2	6	5	NORM	SRV-01	A+
1005	1005-A	DISC	HOUSE	D23	V23	DENV	80-110K	VOD	2013-09-01	PL	4	15	12	NORM	SRV-01	A+

FIG. 11 - Sample Data – Continued

Example 5:

Sample Data for Linear Viewing Activity, DVR, VOD, EDU, LIV

Device Example – Translating Viewing Device Information to a Common Device Type and All Other

Viewer Information – Before

PRO-GRAM	PROGRAM AIRING	CHAN-NEL INFO	HOUSE	VIEWING DEVICE	VIEWER	GEO-GRAPHIC INFO	DEMO-GRAPHIC INFO	VIEWING TYPE	VIEWING DATE	PLAY-BACK MODE	PLAY-BACK BEGIN POSITION	PLAY-BACK END POSITION	ACTIVITY DURATION SECONDS	PRESENT-ATION INFO	VIDEO SERVER INFO	SYSTEM HEALTH INFO
1010	1020	1030	1040	1050	1060	1070	1080	1090	1100	1110	1120	1130	1140	1150	1160	1170
1001	1001-A	HIST	H20	STB	V21	DENV	40-60k	LIN	2013-08-31	PL	3	7	5	NORM	SRV-01	A+
1002	1002-A	HISTHD	H21	IPTV	V21	DENV	40-60k	LIN	2013-08-31	PL	2	6	5	NORM	SRV-01	A+
1003	1003-A	FOX	H23	TABLET	V23	DENV	40-60k	LIN	2013-08-31	PL	3	7	5	NORM	SRV-01	A+
1004	1004-A	FOXHD	H24	PHONE	V24	DENV	40-60k	LIN	2013-08-31	PL	2	6	5	NORM	SRV-01	A+
1005	1005-A	DISC	H25	STB	V23	DENV	80-110K	LIN	2013-09-01	PL	4	15	12	NORM	SRV-01	A+

Viewer Information – After

PRO-GRAM	PROGRAM AIRING	CHAN-NEL INFO	HOUSE	VIEWING DEVICE	VIEWER	GEO-GRAPHIC INFO	DEMO-GRAPHIC INFO	VIEWING TYPE	VIEWING DATE	PLAY-BACK MODE	PLAY-BACK BEGIN POSITION	PLAY-BACK END POSITION	ACTIVITY DURATION SECONDS	PRESENT-ATION INFO	VIDEO SERVER INFO	SYSTEM HEALTH INFO
1210	1220	1230	1240	1250	1260	1270	1280	1290	1300	1310	1320	1330	1340	1350	1360	1370
1001	1001-A	HIST	H20	STB	V21	DENV	40-60k	LIN	2013-08-31	PL	3	7	5	NORM	SRV-01	A+
1002	1002-A	HISTHD	H21	OTHER	V21	DENV	40-60k	LIN	2013-08-31	PL	2	6	5	NORM	SRV-01	A+
1003	1003-A	FOX	H23	OTHER	V23	DENV	40-60k	LIN	2013-08-31	PL	3	7	5	NORM	SRV-01	A+
1004	1004-A	FOXHD	H24	OTHER	V24	DENV	40-60k	LIN	2013-08-31	PL	2	6	5	NORM	SRV-01	A+
1005	1005-A	DISC	H25	STB	V23	DENV	80-110K	LIN	2013-09-01	PL	4	15	12	NORM	SRV-01	A+

FIG. 11 - Sample Data – Continued

Example 6:

Sample Data for Linear Viewing Activity, DVR, VOD, EDU, LIV

Viewer Example – Translating Viewer Information from diverse values to a common value and all other

Before

PRO-GRAM	PROGRAM-AIRING	CHAN-NEL INFO	HOUSE	VIEWING DEVICE	VIEWER	GEO-GRAPHIC INFO	DEMO-GRAPHIC INFO	VIEWING TYPE	VIEWING DATE	PLAY-BACK MODE	PLAY-BACK BEGIN POSITION	PLAY-BACK END POSITION	ACTIVITY DURATION (SECONDS)	PARENT ACTION INFO	VIDEO SERVER INFO	SYSTEM HEALTH INFO
1010	1020	1030	1040	1050	1060	1070	1080	1090	1100	1110	1120	1130	1140	1150	1160	1170
1001	1001-A	HIST	HOUSE	STB	PARENT	DENV	40-60k	LIN	2013-08-31	PL	3	7	5	NORM	SRV-01	A+
1002	1002-A	HISTHD	APT	IPTV	CHILD	DENV	40-60k	LIN	2013-08-31	PL	2	6	5	NORM	SRV-01	A+
1003	1003-A	FOX	HOUSE	TABLET	STUDENT	DENV	40-60k	LIN	2013-08-31	PL	3	7	5	NORM	SRV-01	A+
1004	1004-A	FOXHD	HOUSE	PHONE	PARENT	DENV	40-60k	LIN	2013-08-31	PL	2	6	5	NORM	SRV-01	A+
1005	1005-A	DISC	HOUSE	STB	CHILD	DENV	80-110K	LIN	2013-09-01	PL	4	15	12	NORM	SRV-01	A+

After

PRO-GRAM	PROGRAM-AIRING	CHAN-NEL INFO	HOUSE	VIEWING DEVICE	VIEWER	GEO-GRAPHIC INFO	DEMO-GRAPHIC INFO	VIEWING TYPE	VIEWING DATE	PLAY-BACK MODE	PLAY-BACK BEGIN POSITION	PLAY-BACK END POSITION	ACTIVITY DURATION (SECONDS)	PARENT ACTION INFO	VIDEO SERVER INFO	SYSTEM HEALTH INFO
1210	1220	1230	1240	1250	1260	1270	1280	1290	1300	1310	1320	1330	1340	1350	1360	1370
1001	1001-A	HIST	HOUSE	STB	PARENT	DENV	40-60k	LIN	2013-08-31	PL	3	7	5	NORM	SRV-01	A+
1002	1002-A	HISTHD	APT	IPTV	OTHER	DENV	40-60k	LIN	2013-08-31	PL	2	6	5	NORM	SRV-01	A+
1003	1003-A	FOX	HOUSE	TABLET	OTHER	DENV	40-60k	LIN	2013-08-31	PL	3	7	5	NORM	SRV-01	A+
1004	1004-A	FOXHD	HOUSE	PHONE	PARENT	DENV	40-60k	LIN	2013-08-31	PL	2	6	5	NORM	SRV-01	A+
1005	1005-A	DISC	HOUSE	STB	OTHER	DENV	80-110K	LIN	2013-09-01	PL	4	15	12	NORM	SRV-01	A+

FIG. 11 - Sample Data – Continued

Example 7:

Sample Data for Linear Viewing Activity, DVR, VOD, EDU, LIV

Geographic Example – Translating Geographic Information from a detailed value to a higher value

Geographic Information – Before

PRO-GRAM	PROGRAM AIRING	CHAN-NEL INFO	HOUSE	VIEWING DEVICE	VIEWER	GEO-GRAPHIC INFO	DEMO-GRAPHIC INFO	VIEWING TYPE	VIEWING DATE	PLAY-BACK MODE	PLAY-BACK BEGIN POSITION	PLAY-BACK END POSITION	ACTIVITY DURATION SECONDS	PRESANT-ACTION INFO	VIDEO SERVER INFO	SYSTEM HEALTH INFO
1010	1020	1030	1040	1050	1060	1070	1080	1090	1100	1110	1120	1130	1140	1150	1160	1170
1001	1001-A	HIST	H20	D20	V21	80111	40-60K	EDU	2013-08-31	PL	3	7	5	NORM	SRV-01	A+
1002	1002-A	HISTHD	H21	D21	V21	80112	40-60K	EDU	2013-08-31	PL	2	6	5	NORM	SRV-01	A+
1003	1003-A	FOX	H23	D23	V23	80113	40-60K	EDU	2013-08-31	PL	3	7	5	NORM	SRV-01	A+
1004	1004-A	FOXHD	H24	D24	V24	80114	40-60K	EDU	2013-08-31	PL	2	6	5	NORM	SRV-01	A+
1005	1005-A	DISC	H25	D23	V23	80115	80-110K	EDU	2013-09-01	PL	4	15	12	NORM	SRV-01	A+

Geographic Information – After

PRO-GRAM	PROGRAM AIRING	CHAN-NEL INFO	HOUSE	VIEWING DEVICE	VIEWER	GEO-GRAPHIC INFO	DEMO-GRAPHIC INFO	VIEWING TYPE	VIEWING DATE	PLAY-BACK MODE	PLAY-BACK BEGIN POSITION	PLAY-BACK END POSITION	ACTIVITY DURATION SECONDS	PRESANT-ACTION INFO	VIDEO SERVER INFO	SYSTEM HEALTH INFO
1210	1220	1230	1240	1250	1260	1270	1280	1290	1300	1310	1320	1330	1340	1350	1360	1370
1001	1001-A	HIST	H20	D20	V21	801XX	40-60K	EDU	2013-08-31	PL	3	7	5	NORM	SRV-01	A+
1002	1002-A	HISTHD	H21	D21	V21	801XX	40-60K	EDU	2013-08-31	PL	2	6	5	NORM	SRV-01	A+
1003	1003-A	FOX	H23	D23	V23	801XX	40-60K	EDU	2013-08-31	PL	3	7	5	NORM	SRV-01	A+
1004	1004-A	FOXHD	H24	D24	V24	801XX	40-60K	EDU	2013-08-31	PL	2	6	5	NORM	SRV-01	A+
1005	1005-A	DISC	H25	D23	V23	801XX	80-110K	EDU	2013-09-01	PL	4	15	12	NORM	SRV-01	A+

FIG. 11 - Sample Data – Continued

Example 8:

Sample Data for Linear Viewing Activity, DVR, VOD, EDU, LIV

Demographic Example – Translating Demographic Information from a detailed value to a higher value

Demographic Information – Before

PRO-GRAM	PROGRAM AIRING	CHAN-NEL INFO	HOUSE	VIEWING DEVICE	VISWER	GEO-GRAPHIC INFO	DEMO-GRAPHIC INFO	VIEWING TYPE	VIEWING DATE	PLAY-BACK MODE	PLAY-BACK BEGIN POSITION	PLAY-BACK END POSITION	ACTIVITY DURATION SECONDS	PRESET-SECTION INFO	VIDEO SERVER INFO	SYSTEM SERIAL# INFO
1010	1020	1030	1040	1050	1060	1070	1080	1090	1100	1110	1120	1130	1140	1150	1160	1170
1001	1001-A	HIST	H20	D20	V21	DENV	1-3 YR	LIV	2013-08-31	PL	3	7	5	NORM	SRV-01	A+
1002	1002-A	HISTHD	H21	D21	V21	DENV	4-5 YR	LIV	2013-08-31	PL	2	6	5	NORM	SRV-01	A+
1003	1003-A	FOX	H23	D23	V23	DENV	6-8 YR	LIV	2013-08-31	PL	3	7	5	NORM	SRV-01	A+
1004	1004-A	FOXHD	H24	D24	V24	DENV	4-5 YR	LIV	2013-08-31	PL	2	6	5	NORM	SRV-01	A+
1005	1005-A	DISC	H25	D23	V23	DENV	21-44 YR	LIV	2013-09-01	PL	4	15	12	NORM	SRV-01	A+
1006	1006-A	ABC	H26	D26	V26	DENV	45-54 YR	LIV	2013-09-01	PL	4	15	12	NORM	SRV-01	A+

Demographic Information – After

PRO-GRAM	PROGRAM AIRING	CHAN-NEL INFO	HOUSE	VIEWING DEVICE	VISWER	GEO-GRAPHIC INFO	DEMO-GRAPHIC INFO	VIEWING TYPE	VIEWING DATE	PLAY-BACK MODE	PLAY-BACK BEGIN POSITION	PLAY-BACK END POSITION	ACTIVITY DURATION SECONDS	PRESET-SECTION INFO	VIDEO SERVER INFO	SYSTEM SERIAL# INFO
1210	1220	1230	1240	1250	1260	1270	1280	1290	1300	1310	1320	1330	1340	1350	1360	1370
1001	1001-A	HIST	H20	D20	V21	DENV	CHILD	LIV	2013-08-31	PL	3	7	5	NORM	SRV-01	A+
1002	1002-A	HISTHD	H21	D21	V21	DENV	CHILD	LIV	2013-08-31	PL	2	6	5	NORM	SRV-01	A+
1003	1003-A	FOX	H23	D23	V23	DENV	CHILD	LIV	2013-08-31	PL	3	7	5	NORM	SRV-01	A+
1004	1004-A	FOXHD	H24	D24	V24	DENV	CHILD	LIV	2013-08-31	PL	2	6	5	NORM	SRV-01	A+
1005	1005-A	DISC	H25	D23	V23	DENV	ADULT	LIV	2013-09-01	PL	4	15	12	NORM	SRV-01	A+
1006	1006-A	ABC	H26	D26	V26	DENV	ADULT	LIV	2013-09-01	PL	4	15	12	NORM	SRV-01	A+



**FIG. 12**  
**Summary Information**

Specification Name:	Prepared Video Viewing Activity Data File <b>160</b>
Record description:	<p>The Prepared Video Viewing Activity Data File contains video viewing activity records resulting from either the Channel Translation Processor <b>150</b> or the Data Translation Processor <b>154</b> having processed the Video Viewing Activity Data File <b>130</b> and applying various translation rules to the file in order to create the Prepared Video Viewing Activity Data File which is ready for downstream processing.</p> <p>These prepared records may be based on upon Linear Tuning Activity (LTA), sessionized linear viewing activity (LVA), digital video recording viewing activity (DVR), or video on demand viewing activity (VOD), Educational viewing activity (EDU), and Live (LIV) viewing. The Prepared Video Viewing Activity Data File <b>160</b> can hold any of these formats with the various fields being populated as needed.</p>

**Data Structure**

1200	1202	1204	1206
NAME	TYPE	SIZE	DESCRIPTION
<p><b>Notes:</b> The <b>CHANNEL_INFO_TARGET 1400</b> field is shown below. This allows the analyst to see both the original Channel information and the enriched Channel information.</p> <p>In <b>FIG 3</b> I have shown the translated Channel value being appended to the file as field <b>CHANNEL_INFO_TARGET 1400</b> as shown here in <b>FIG 12</b>.</p> <p>In <b>FIG 4</b> I have shown that the value which is being replaced can be loaded to a field with the same name in the prepared file as in the source file such as is shown here in <b>FIG 12</b>. In an alternative embodiment, the analyst can define fields for each value that is being replaced so that the prepared (enriched) record layout would contain both the original value and the substituted value in each record. This would require additional space, but has the benefit of showing the original and translated values for audit and any other purpose that may require such values.</p>			

FIG 12 Data Structure - Continued

1210~	PROGRAM__INFO	CHAR	40	The Program Information field. See 1010 for definition.
1220~	PROGRAM_AIRING__INFO	CHAR	40	The Program Airing Information field. See 1020 for definition.
1230~	CHANNEL__INFO	CHAR	40	The Channel Information field. See 1030 for definition.
1240~	HOUSE__INFO	CHAR	40	The House Information field. See 1040 for definition.
1250~	DEVICE__INFO	CHAR	40	The Device Information field. See 1050 for definition.
1260~	VIEWER__INFO	CHAR	40	The Viewer Information field. See 1060 for definition.
1270~	GEOGRAPHIC_INF O	CHAR	40	The Geographic Information field. See 1070 for definition.
1280~	DEMOGRAPHIC_IN FO	CHAR	40	The Demographic Information field. See 1080 for definition.
1290~	VIEWING__TYPE	CHAR	3	The Viewing Type. See 1090 for definition.
1300~	VIEWING__DATE	DATE	8	The Viewing Date. See 1100 for definition.
1302~	TUNE_IN_DATE__T IME	DATE TIME		The Tune in date time. See 1102 for definition.

FIG 12 Data Structure - Continued

1304~	TUNE_IN_SECOND__OF_DAY	Number	5	Tune In Second of Day. See 1104 for definition.
1306~	TUNE_OUT_DATE_TIME	DATE TIME		Tune out date and time. See 1106 for definition.
1308~	TUNE_OUT_SECONDS_OF_DAY	Number	5	Tune out Second of Day. See 1108 for definition.
1310~	PLAYBACK_MODE	CHAR	2	The Playback Mode. See 1110 for definition.
1320~	PLAYBACK_BEGIN_POSITION	Number	5	The Playback Begin Position. See 1120 for definition.
1330~	PLAYBACK_END_POSITION	Number	5	The Playback End Position. See 1130 for definition.
1338~	TUNE_DURATION_SECONDS	Number	5	Tune Duration Seconds. See 1138 for definition.
1340~	ACTIVITY_DURATION_SECONDS	Number	5	The Activity Duration. See 1140 for definition.
1350~	PRESENTATION_INFO	CHAR	4	The Presentation Information. See 1150 for definition.
1360~	VIDEO_SERVER_INFO	CHAR	40	The Video Server Information. See 1160 for definition.
1370~	SYSTEM_HEALTH_INFO	CHAR	40	The System Health Information. See 1170 for definition.
1400~	CHANNEL_INFORMATION_TARGET	CHAR	40	Channel Information Target shows the translated value assigned to the channel information. As noted above, the translated value can be carried in the existing field or a new field can be added as shown here.

## DATA TRANSLATION FOR VIDEO-VIEWING ACTIVITY

### CROSS-REFERENCE TO RELATED APPLICATIONS

**[0001]** This application is related to U.S. Pat. No. 8,365,212 B1 issued on Jan. 29, 2013 entitled “SYSTEM AND METHOD FOR ANALYZING HUMAN INTERACTION WITH ELECTRONIC DEVICES THAT ACCESS A COMPUTER SYSTEM THROUGH A NETWORK” by the present inventor which is incorporated by reference in its entirety but is not admitted to be prior art.

**[0002]** This application is also related to U.S. Pat. No. 8,365,213 B1 issued on Jan. 29, 2013 entitled “SYSTEM AND METHOD FOR MEASURING TELEVISION ADVERTISING AND PROGRAM VIEWING AT A SECOND-BY-SECOND LEVEL AND FOR MEASURING EFFECTIVENESS OF TARGETED ADVERTISING” by the present inventor which is incorporated by reference in its entirety but is not admitted to be prior art.

**[0003]** This application is also related to the co-pending application U.S. application Ser. No. 13/360,704 filed on Jan. 28, 2012 entitled “SYSTEM AND METHOD FOR MEASURING LONGITUDINAL VIDEO ASSET VIEWING AT A SECOND-BY-SECOND LEVEL TO UNDERSTAND BEHAVIOR OF VIEWERS AS THEY INTERACT WITH VIDEO ASSET VIEWING DEVICES THAT ACCESS A COMPUTER SYSTEM THROUGH A NETWORK” by the present inventor which is incorporated by reference in its entirety but is not admitted to be prior art.

**[0004]** This application is also related to the co-pending application U.S. application Ser. No. 13/567,073 filed on Aug. 5, 2012 entitled “SYSTEM AND METHOD FOR MEASURING LINEAR, DVR, AND VOD VIDEO PROGRAM VIEWING AT A SECOND-BY-SECOND LEVEL TO UNDERSTAND BEHAVIOR OF VIEWERS AS THEY INTERACT WITH VIDEO ASSET VIEWING DEVICES DELIVERING CONTENT THROUGH A NETWORK” by the present inventor which is incorporated by reference in its entirety but is not admitted to be prior art.

**[0005]** This application is also related to the co-pending application U.S. application Ser. No. 13/740,199 filed on Jan. 13, 2013 entitled “SYSTEM AND METHOD FOR MEASURING DEMOGRAPHIC-BASED HOUSEHOLD ADVERTISING REACH; IMPRESSIONS, SHARE, HUT, RATING, AND CUMULATIVE AUDIENCE; AND VIDEO PROGRAM VIEWING, BASED ON SECOND-BY-SECOND HOUSE LEVEL VIEWING ACTIVITY, TO UNDERSTAND BEHAVIOR OF VIEWERS AS THEY INTERACT WITH VIDEO ASSET VIEWING DEVICES DELIVERING CONTENT THROUGH A NETWORK” by the present inventor which is incorporated by reference in its entirety but is not admitted to be prior art.

**[0006]** This application is also related to the co-pending application U.S. application Ser. No. 14/013,031 filed on Aug. 29, 2013 entitled “SYSTEM AND METHOD FOR USING THE HADOOP MAPREDUCE FRAMEWORK TO MEASURE VIDEO CONTENT VIEWING ON SECOND-BY-SECOND LEVEL TO UNDERSTAND BEHAVIOR OF VIEWERS AS THEY INTERACT WITH VIDEO ASSET VIEWING DEVICES DELIVERING CONTENT THROUGH A NETWORK” by the present inventor which is incorporated by reference in its entirety but is not admitted to be prior art.

**[0007]** This application is also related to the co-pending application U.S. application Ser. No. 14/020,778 filed on Sep. 6, 2013 entitled “SYSTEM AND METHOD FOR USING THE HADOOP MAPREDUCE FRAMEWORK TO MEASURE LINEAR, DVR, AND VOD VIDEO PROGRAM VIEWING INCLUDING MEASURING TRICK PLAY ACTIVITY ON SECOND-BY-SECOND LEVEL TO UNDERSTAND BEHAVIOR OF VIEWERS AS THEY INTERACT WITH VIDEO ASSET VIEWING DEVICES DELIVERING CONTENT THROUGH A NETWORK” by the present inventor which is incorporated by reference in its entirety but is not admitted to be prior art.

**[0008]** A portion of the disclosure of this patent document contains material which is subject to copyright protection. The copyright owner has no objection to the facsimile reproduction by anyone of the patent document or the patent disclosure, as it appears in the Patent and Trademark Office patent file or records, but otherwise reserves all copyright rights whatsoever.

### BACKGROUND

**[0009]** 1. Prior Art

**[0010]** I have not found any relevant prior art at the present time.

**[0011]** 2. Background Information

#### General Statement of Problem

**[0012]** With the ever increasing number of consumer choices for television viewing, it is important for advertisers, content producers, and service providers such as cable television and satellite television and internet protocol television companies to be able to accurately measure audience viewership. I have discussed this problem extensively in my prior applications. In two of my previous Applications, I have taught how to use the MapReduce Framework to aggregate video viewing activity. When using that framework an analyst may find that certain aggregation functions consume a large quantity of computing resources but do not add value commensurate with that consumption of resources. This application teaches how to reduce the workload on the MapReduce Framework by implementing various data translation strategies prior to sending the video viewing activity file to downstream processes for measuring or aggregating Linear, DVR, and VOD Second-By-Second Video Viewing Activity. The reader will readily recognize that these aggregation strategies are applicable to a wide range of video content including, but not limited to, cable and/or satellite television video on demand, all variety of video content delivered electronically across a network, and online educational videos.

**[0013]** Existing Tools for Data Analysis

**[0014]** In my prior applications U.S. application Ser. No. 14/013,031 filed on Aug. 29, 2013 and U.S. application Ser. No. 14/020,778 filed on Sep. 6, 2013 I taught how to aggregate video viewing activity and video program viewing activity using the MapReduce Framework. I have not found any other teachings on this topic.

### SUMMARY

**[0015]** In accordance with one embodiment, I disclose a computer-implemented method of using Linear, DVR, and VOD video viewing activity data as input to a data translation processor which prepares that video viewing activity for more efficient downstream processing by translating detailed val-

ues to aggregated values according to analyst defined translation rules in preparation for ingestion by a MapReduce Framework with the result that the MapReduce Framework needs to process less data in order to create analytical studies of second-by-second viewing activity for program, channel, house, device, viewer, demographic, and geographic attributes in combination or as individual attributes. Once the data translation rules have been applied to the video viewing activity by the data translation processor, the data is ready for use by downstream tools such as the MapReduce Framework which is able to aggregate the data much more efficiently than would have been possible prior to running the data translation processor. Additionally, by applying the translations to the video viewing activity file which contains all of the detailed values, this enables the analyst to use a single copy of that file for multiple analytical studies thus avoiding the time and cost associated with creating a new extract file with embedded translated values for each analytical study.

#### Advantages

**[0016]** By implementing the data translation processor that I teach about in this Application, an analyst can produce video viewing aggregations using the MapReduce distributed computing framework in less time using less computing resources. This will allow the analyst to aggregate larger data sets than would otherwise be possible using a given set of computer hardware. It will also allow the analyst to run additional studies thus potentially gaining additional insights into viewing behaviors.

**[0017]** Additionally, the analyst is able to save significant time and computing resources by avoiding the need to recreate the video viewing activity file with multiple data translations embedded in it because he does not need a new extract from the source system for each analytical study.

#### DEFINITIONS

**[0018]** The following are definitions that will aid in understanding one or more of the embodiments presented herein:

**[0019]** Computer readable format means any data format that can be read by a computer program or a human being as necessary. Nonlimiting examples include:

**[0020]** (i) formatted text files,

**[0021]** (ii) pipe delimited text files,

**[0022]** (iii) data base tables,

**[0023]** (iv) Extensible Markup Language (XML) messages,

**[0024]** (v) a printed report,

**[0025]** (vi) JavaScript Object Notation messages.

**[0026]** Data analysis computer system means a combination of one or more computers on which a Data Translation Program or Programs can be executed. Nonlimiting examples include:

**[0027]** (i) one or more computers where video viewing activity data can be used as input to a process which creates prepared video viewing activity data.

**[0028]** Data analysis computer of known type means any commonly available computer system running a commonly known operating system. Nonlimiting examples include:

**[0029]** (i) a standard personal computer running WINDOWS 7 Professional operating system from MICROSOFT® Corporation,

**[0030]** (ii) a computer running the UNIX operating system,

**[0031]** (iii) a computer running the Linux operating system,

**[0032]** (iv) a computer in a cloud computing environment,

**[0033]** (v) a mainframe computer with its operating system.

**[0034]** Data translation program means a computer program or programs that are able to execute on a Data analysis computer of known type. Nonlimiting examples include:

**[0035]** (i) a JAVA program, (ii) a Python script, (iii) a COBOL program.

**[0036]** Demographic information means any data item that can describe a characteristic of a viewer or a subscriber or a household associated with a viewer who is operating the video asset viewing device. Nonlimiting examples include income, ethnicity, gender, age, marital status, location, geographic area, postal code, census data, occupation, social grouping, family status, any proprietary demographic grouping, segmentation, credit score, dwelling type, homeownership status, property ownership status, rental status, vehicle ownership, tax rolls, credit card usage, religious affiliation, sports interest, political party affiliation, cable television subscriber type, cable television subscriber package level, and cell phone service level.

**[0037]** Device Characteristic means any feature or capability or aspect or descriptive qualifier or identifier of a video viewing device. Nonlimiting examples include that this may identify the type of device such as a set-top box, a tablet, a smart phone; a capability of the device such as the ability to record video or to support multiple viewing windows, or a manufacturer identifier.

**[0038]** Device Type is a subset of Device Characteristic where device type may, as a nonlimiting example, identify the type of device such as a set-top box, a tablet, a smart phone.

**[0039]** Digital Video Recorder means a device that records video content from a network for later playback. This includes but is not limited to set-top box DVR, network DVR, and cloud DVR.

**[0040]** DVR—see Digital Video Recorder.

**[0041]** Digital Video Recording (DVR) Playback is when the viewer plays back content that was previously recorded on their DVR. DVR content can be viewed using various Trick Play features.

**[0042]** Geographic information means any service area or any network hierarchy designation or marketing area or other designated area used by a cable television company or a satellite television company or IP Television delivery company or educational service provider or video asset delivery system. The boundary or description of a geographic area is defined based on the needs of the service provider. Nonlimiting examples include a Market in a cable company network, a Headend in a cable company network, a Hub in a cable company network, a census tract, a cell tower identifier, a service area for satellite TV, advertising zone, a zip code, or some other geographic identifier. The geographic information may then be used to identify the location of a video asset viewing device or geographic information about the about the house associated with the device or the location of the device at the time of the viewer interaction in the event that the viewer interaction occurs in a location different than the location of the house associated with the device.

**[0043]** Network means any computer network. Nonlimiting examples include:

**[0044]** (i) a cable television network,

**[0045]** (ii) a cellular telephony network,

**[0046]** (iii) hybrid fiber coax system,

**[0047]** (iv) a satellite television network,

**[0048]** (v) a wi-fi network,

**[0049]** (vi) any means that supports communication among video asset viewing devices or electronic devices or computers or computer systems.

**[0050]** Pipe delimited text files means data files where the fields are separated by the “|” character.

**[0051]** New form of said video viewing activity data means the prepared version of the Video Viewing Activity Data File.

**[0052]** Sessionized Linear Viewing is linear tuning activity which has been transformed into program based viewing. A simple linear tuning event may cross program boundaries. That simple linear tuning event can be split into multiple program based linear viewing activity records by creating separate tuning records for each program that is viewed during the linear tuning event. The viewer may use “trick plays” when viewing this content. Additionally, the sessionized linear viewing may represent live viewing activity or time shifted viewing activity.

**[0053]** Set-top box means a video asset viewing device that receives external signals and decodes those signals into content that can be viewed on a television screen or similar display device. The signals may come from a cable television system, a satellite television system, a network, or any other suitable means. A set-top box may have one or more tuners. The set-top box allows the user to interact with it to control what is displayed on the television screen. The set-top box is able to capture the commands given by the user and then transmit those commands to another computer system. For purposes of this application, stating that a set-top box tunes to a channel is equivalent to stating that a tuner in a set-top box has tuned to a channel. A set-top box may also play back previously recorded video content.

**[0054]** STB means Set-top box.

**[0055]** Trick Play means using features of the video viewing device to execute operations such as Play, Fast Forward at various speeds (1x, 2x, 3x, 4x), Pause, Skip, Reverse at various speeds (1x, 2x, 3x, 4x), Slow play, slow reverse, and similar activities.

**[0056]** Tuner means a tuner in a Set-top box.

**[0057]** Tuner index means an identifier of a tuner in a Set-top box.

**[0058]** Video On Demand (VOD) a video service whereby previously recorded video content is made available for viewing. VOD content can be viewed using various Trick Play features. The content may include, but is not limited to, cable and/or satellite television video on demand, all variety of video content delivered electronically across a network, and online educational videos.

**[0059]** Video asset means any programming content that may be viewed and/or heard. A Video Program may contain multiple Video Assets. Nonlimiting examples of Video Asset include:

**[0060]** (i) advertisements or commercials, (ii) movies, (iii) sports programs, (iv) news casts, (v) music, (vi) television programs, (vii) video recordings.

**[0061]** Video asset viewing device means any electronic device that may be used either directly or indirectly by a human being to interact with video content where the video content is provided by a cable television system or a satellite television system or a computer system accessed through a network. Nonlimiting examples include: Gaming station, web browser, MP3 Player, Internet Protocol phone, Internet Protocol television, mobile device, mobile smart phone, set-top box, satellite television receiver, set-top box in a cable television network, set-top box in a satellite television sys-

tem, cell phone, personal communication device, personal video recorder, personal video player, two-way interactive service platforms, personal computer, tablet device.

**[0062]** Video server delivering video content through a network means any computer system, any individual piece of computer equipment or electronic gear, or any combination of computer equipment or electronic gear which enables or facilitates the viewer interaction with the video asset viewing device. Nonlimiting examples include:

**[0063]** (i) cable television system, (ii) cable television switched digital video system,

**[0064]** (iii) cellular phone network, (iv) satellite television system, (v) web server,

**[0065]** (vi) any individual piece of computer equipment or electronic gear,

**[0066]** (vii) any combination of computer equipment or electronic gear.

**[0067]** Video viewing activity means any identifiable activity that a Video asset viewing device operator may do in regard to a Video asset viewing device and where such activity can be captured by the video asset viewing device or by the video server delivering video content through a network that supports the device. Nonlimiting examples include:

**[0068]** (i) power on/power off, open web page, close web page,

**[0069]** (ii) channel up/channel down/channel selection, play video content on web browser,

**[0070]** (iii) volume up/volume down/mute/unmute,

**[0071]** (iv) any trick play such as fast forward, rewind, pause

**[0072]** (v) recording video content,

**[0073]** (vi) playing back recorded video content,

**[0074]** (vii) invoking a menu, choosing a menu option,

**[0075]** (viii) any response to a screen prompt

**[0076]** (ix) playing live video content.

**[0077]** Viewer means the human being causing a Viewer interaction; the user of a Set-top box or a Video asset viewing device.

#### BRIEF DESCRIPTION OF THE DRAWINGS

**[0078]** FIG. 1 provides an overview of an exemplary process for collecting viewer interaction data derived from a plurality of viewers interacting with video content that was delivered on a plurality of video asset viewing devices delivering content through a network and then loading that data to a Media Measurement Data Base.

**[0079]** FIG. 2 illustrates an exemplary process for creation of a Video Viewing Activity Data File from a Media Measurement Data Base in preparation for feeding the file into a downstream process.

**[0080]** FIG. 3 illustrates an exemplary process for enriching a Video Viewing Activity Data File with translated channel information in preparation for feeding the file into a downstream process.

**[0081]** FIG. 4 illustrates an exemplary process for enriching a Video Viewing Activity Data File with translations of various fields in preparation for feeding the file into a downstream process.

**[0082]** FIG. 5 illustrates an exemplary record layout for a Video Viewing Activity Data File **130** record, according to one embodiment.

**[0083]** FIG. 6 illustrates an exemplary record layout for a Channel Information Translation Table **140** or **142** record along with sample data, according to one embodiment.

[0084] FIG. 7 illustrates an exemplary data structure design for a Channel Mapping Definition Array Data Structure, according to one embodiment.

[0085] FIG. 8 illustrates an exemplary record layout for a Data Translation Table 146 or 148 record along with sample data, according to one embodiment.

[0086] FIG. 9 illustrates an exemplary data structure design for a Generalized Mapping Definition Array Data Structure, according to one embodiment.

[0087] FIG. 10 illustrates various non-limiting examples which provide 'Before' and 'After' images of Linear Tuning Activity (LTA) records to illustrate what the Channel Translation Processor 150 and Data Translation Processor 154 do, according to one embodiment.

[0088] FIG. 11 illustrates various non-limiting examples which provide 'Before' and 'After' images of Linear Viewing Activity (LVA), Digital Video Recorder viewing activity (DVR), Video On Demand viewing activity (VOD), Educational viewing (EDU), and Live viewing activity (LIV) records to illustrate what the Channel Translation Processor 150 and Data Translation Processor 154 do, according to one embodiment.

[0089] FIG. 12 illustrates an exemplary record layout for a Prepared Video Program Viewing Activity Data File 160 record created by either the Channel Translation Processor 150 or the Data Translation Processor 154, according to one embodiment. There is Summary Information followed by the Data Structure including field definitions.

#### DETAILED DESCRIPTION OF THE DRAWINGS

[0090] When reading the information below, it can be appreciated that these are merely samples of table layouts, format and content, and many aspects of these tables may be varied or expanded within the scope of the embodiment. The table layouts, field formats and content, algorithms, and other aspects are what I presently contemplate for this embodiment, but other table layouts, field formats and content, algorithms, etc. can be used. The algorithms are samples and various aspects of the algorithms may be varied or expanded within the scope of the embodiment.

[0091] In one embodiment the Channel Translation Processor 150 and the Data Translation Processor 154 can be implemented on computer clusters accessing a distributed file system under the Linux operating system. The Channel Translation Processor 150 and the Data Translation Processor 154 can each be implemented in JAVA or Python or COBOL or various other languages. Those skilled in the art will readily recognize these tools.

[0092] Note on Media Measurement Data Model

[0093] Cable Television Laboratories, Inc. has published an "Audience Data Measurement Specification" as "Open-Cable™ Specifications, Audience Measurement, Audience Measurement Data Specification" having Document Control Number "OC-SP-AMD-101-130502" copyright © Cable Television Laboratories, Inc. 2013 which describes a Media Measurement Data Model (MMDM) database design which can be used as a source of data for both Channel Translation Processor 150 and the Data Translation Processor 154 which I teach how to build in this Application. The teaching in my present application can be implemented in conjunction with that Media Measurement Data Model or with any number of data models as long as the required input data is provided as described herein.

[0094] Note:

[0095] Numbering in the Drawings—The numbers in the drawings are usually, but not always, in sequential order.

[0096] FIG. 1 provides an overview of an exemplary process for collecting viewer interaction data derived from a plurality of viewers interacting with video content that was delivered on a plurality of video asset viewing devices delivering content through a network and then loading that data to a Media Measurement Data Base 100. This figure illustrates several viewers interacting with video asset viewing devices to view content which was delivered to those devices across a network and then to collect viewing activity from those devices.

[0097] In this nonlimiting example, the purpose is not to describe in detail the operations of a video content delivery network or a data collection process, but simply to show how the data that is collected from that system can be made available to my Channel Translation Processor 150 or my Data Translation Processor 154.

[0098] It begins with Viewer Viewing Linear Content 9200 who is interacting with a set-top box 9210 and television 9220 as he views linear content. The set-top box 9210 interacts with a Video Content Delivery System 9250 which delivers the content across a Network 9230.

[0099] It continues with Viewer Viewing DVR Content 9202 who is interacting with a set-top box 9210 and television 9220 as he interacts with DVR content, recording content and playing back recorded content using various modes including trick plays. The set-top box 9210 interacts with a Video Content Delivery System 9250 which delivers the content across a Network 9230.

[0100] It continues with Viewer Viewing VOD Content 9203 who is interacting with a set-top box 9210 and television 9220 as he interacts with VOD content, playing the content using various modes including trick plays. The set-top box 9210 interacts with a Video Content Delivery System 9250 which delivers the content across a Network 9230.

[0101] It continues with Viewer viewing video content using tablet, smart phone, IP TV, or other video viewing device 9204 who is interacting with a variety of Video Viewing Devices 9212, including but not limited to tablet, smart phone, IP TV, PC, etc. The video viewing device interacts with a Video Content Delivery System 9250 which delivers the content across a Network 9230.

[0102] Video Content Delivery System 9250 then interacts with a Viewer Interaction Data, Data Collection System 9260 which collects all manner of viewer interaction data including Linear viewing including time-shifted linear viewing, Digital Video Recorder recording and playback/viewing, Video on Demand viewing, Educational video viewing, Streaming video viewing, and Live viewing. The Viewer Interaction Data, Data Collection System 9260 then processes the data as needed to load it to a Media Measurement Data Base 100. The data in the Media Measurement Data Base 100 can then be used as input to creating the Video Viewing Activity Data File 130, as described in FIG. 2.

[0103] FIG. 2 illustrates an overview of the process for creating the Video Viewing Activity Data File 130. The process of extracting this file from the Media Measurement Database 100 is often an expensive operation. For this reason, an analyst may wish to create this file in a somewhat generalized format with numerous detail fields and then store the file on a distributed file system in preparation for running a variety of analytical processes against the file.

**[0104]** As a non-limiting example, a file of Video Viewing Activity Data may be used to analyze viewing in numerous dimensions such as:

- [0105]** (i) program information, (ii) channel information, (iii) house information,
- [0106]** (iv) device information, (v) viewer information, (vi) geographic information,
- [0107]** (vii) demographic information.

**[0108]** To provide a number of non-limiting examples, for program information, the analyst may need to:

- [0109]** a) Analyze sports viewing vs. all other viewing
- [0110]** b) Analyze News viewing vs. sports viewing
- [0111]** c) Analyze Sports viewing vs. movie viewing
- [0112]** d) Analyze "G" program viewing vs. all other viewing.

**[0113]** To provide a number of non-limiting examples, for channel information, the analyst may need to:

- [0114]** a) Analyze standard definition viewing vs. high definition viewing of a channel
- [0115]** b) Analyze all standard definition viewing vs. all high definition viewing
- [0116]** c) Combine standard definition viewing and high definition viewing to measure that
- [0117]** d) Analyze combined standard definition viewing and high definition viewing of one channel vs. all other channels.

**[0118]** To provide a number of non-limiting examples, for house information, the analyst may need to:

- [0119]** a) Analyze residential viewing vs. commercial viewing
- [0120]** b) Analyze house viewing vs. apartment (multi-family) viewing.

**[0121]** To provide a number of non-limiting examples, for device information, the analyst may need to:

- [0122]** a) Analyze STB viewing vs. all other viewing
- [0123]** b) Analyze STB viewing vs. tablet viewing.

**[0124]** To provide a number of non-limiting examples, for viewer information, the analyst may need to:

- [0125]** a) Analyze viewing by viewer type such as parent vs. student
- [0126]** b) Analyze viewing by grade level.

**[0127]** To provide a number of non-limiting examples, for geographic information, the analyst may need to:

- [0128]** a) Analyze viewing by a combination of zip codes
- [0129]** b) Analyze viewing by one marketing area compared to another marketing area.

**[0130]** To provide a number of non-limiting examples, for demographic information, the analyst may need to:

- [0131]** a) Analyze viewing by a combination of age levels
- [0132]** b) Analyze viewing by one education level vs. another education level
- [0133]** c) Analyze viewing by one education level vs. all other education levels.

**[0134]** In addition to all of these possibilities, an analyst may combine any number of these qualifiers.

**[0135]** To provide a number of non-limiting examples, for demographic information, the analyst may need to:

- [0136]** a) Analyze viewing of a channel by a combination of age levels
- [0137]** b) Analyze viewing of a channel by device types
- [0138]** c) Analyze viewing by a program by demographic groups

**[0139]** d) Analyze viewing by a program by demographic groups by device types.

**[0140]** In each of the examples above, the analyst may be able to take the same Video Viewing Activity Data File **130** that was extracted from the Media Measurement Database **100** and use it, with various enrichments, to feed into a MapReduce process where he can aggregate the data using hundreds or even thousands of computers working in parallel. Those skilled in the art will readily recognize that when there are fewer unique keys to aggregate using MapReduce, the MapReduce process will consume fewer resources, thus allowing the analyst to complete more studies in less time with less computing resources.

**[0141]** To explain this further, in my pending applications U.S. application Ser. No. 14/013,031 filed on Aug. 29, 2013 and U.S. application Ser. No. 14/020,778 filed on Sep. 6, 2013 I taught how to use the MapReduce distributed computing framework to analyze video viewing activity. In those applications I taught how to explode the video viewing activity so that one record is created for each second of viewing activity for each incoming record.

**[0142]** If an analyst was analyzing 3600 seconds of viewing across 100 channels, this could potentially lead to  $3600 \times 100 = 360,000$  unique keys in the Reduce part of the MapReduce process. However, if the analyst knew in advance that he was only interested in detailed viewing information for 10 channels and all the other activity could be grouped into an "OTHER" bucket, then this could potentially lead to  $3600 \times 11 = 39,600$  unique keys in the Reduce part of the MapReduce process. This is a very large reduction in the unique key count which results in a significant reduction in run time and computing resources needed to run the analysis.

**[0143]** Similarly, suppose there is a cable system with 30° channels. If an analyst was interested in analyzing the viewing activity of 20 channel across the entire day of 86,400 seconds, he would likely still need to measure the viewing activity of all the other 280 channels, perhaps to be able to calculate the percentage of the overall viewing activity at any second of the day for each of the 20 channels. In this case assume the analyst is not interested in the activity of those other channels except to have an accurate count of total viewing during any second of the day. If the analyst simply aggregated the incoming data as it is, the result could potentially be  $300 \times 86,400 = 25,920,000$  unique keys in the Reduce part of the MapReduce process. By grouping the viewing activity of the 280 channels into an "OTHER" bucket, then this could potentially lead to  $86,400 \times 21 = 1,814,400$  unique keys in the Reduce part of the MapReduce process. This is a very large reduction in the unique key count which results in a significant reduction in run time and computing resources needed to run the analysis.

**[0144]** This same thought process is application to program information, house information, device information, viewer information, geographic information, and demographic information. Thus we see that the same extract file can be used in a multitude of analytical processes if the analyst is able to apply some pre-aggregation rules to the file to create a Prepared Video Viewing Activity Data File. In the remainder of this specification I will teach how to create such an Prepared Video Viewing Activity Data File which can then be provided to downstream analytic processes such as the Data Explosion Process described in my previous Applications.

**[0145]** Before resuming the review of FIG. 2, note that the video viewing activity may be sourced from a Media Mea-



surement Database such as the one described in the Cable Television Laboratories, Inc. specification. The populating of the Media Measurement Database **100** is beyond the scope of this application and so only brief remarks will be made in reference to that. There are video viewing data collection systems that are commonly used in the industry for collecting channel tuning or video viewing activity data including switched digital video systems, set top box applications, internet protocol video viewing applications, and other video viewing applications. I have discussed these in detail in my prior applications. These systems enable the collection of the video viewing events which can be loaded to a Media Measurement Database **100**. From such a database, Video Viewing Activity Data can be extracted in a format similar to that shown in FIG. 5 Video Viewing Activity Data File **130**.

[0146] Also before resuming the review of FIG. 2, it is helpful to recognize that creating the Video Viewing Activity Data file is often a time consuming operation. It can be one of the slower tasks in the entire chain of events in analyzing the video viewing activity. For this reason, it is helpful to avoid multiple or repeated data extraction processes. One solution to this problem is to extract a single file which contains all of the video viewing activity in a detailed format (with numerous fields), perhaps on a daily basis for the previous day's viewing activity, and then save that detailed file on a distributed file system. Then that detail file can be used as input to a translation process such as I describe herein where the detailed values are translated to or replaced by summary values as needed to perform the analytical study of interest. By designing the system in this way, the analyst can take advantage of the power of massively parallel systems to create analytical studies very quickly.

[0147] Proceeding with the review of FIG. 2, the process begins with Video Viewing Activity File Extract Process Overview **110**. The first step is to extract the video viewing events as per Extract Video Viewing Activity Data from Media Measurement Data Base **120**. Those skilled in the art will have no difficulty creating a database query or similar process to extract data from a Media Measurement Database **100** or other source and making it available in a format similar to that defined in Video Viewing Activity Data File **130**. The file structure is defined in FIG. 5 Video Viewing Activity Data File **130** which describes an exemplary format for the input video viewing activity data.

[0148] This data file may contain various types of viewing activity depending on the query defined by the analyst. A non-limiting example of the types are:

- [0149] a) Linear tuning activity (LTA)
- [0150] b) Sessionized linear viewing activity (LVA)
- [0151] c) Digital video recording viewing activity (DVR)
- [0152] d) Video on demand viewing activity (VOD)
- [0153] e) Education video viewing activity (EDU)
- [0154] f) Live viewing activity (LIV).

[0155] These are defined next:

[0156] Linear tuning activity (LTA)—Linear Tuning Activity is tuning activity which is based on linear tuning events where the viewer tunes to a channel, stays on the channel for some period of time (one second to multiple hours), and then tunes away by tuning to another channel or by initiating some other activity. Linear Tuning Activity may cross program boundaries.

[0157] Sessionized linear viewing activity (LVA)—Sessionized linear viewing activity may be derived from linear

tuning events or it may be captured by the set top box as sessionized activity. For linear tuning events which fit within program boundaries, the tuning event may be enriched as described in the Cable Television Laboratories, Inc. specification. For linear tuning events which cross program boundaries, the tuning event is divided into shorter duration tuning events as needed to fit within program boundaries as described in the specification; the tuning events are also enriched with additional fields. Because the viewing activity has been transformed to fit within program boundaries, I refer to it as program-based viewing. Linear viewing activity may be extracted from the Cable Television Laboratories, Inc. MMDM or from any source that is able to provide the data in a format suitable for this process.

[0158] A non-limiting example will help to explain this:

[0159] Device **100** has a linear tuning event consisting of a Tune to ABC at 6:55:00 PM and tune away from ABC at 8:07:59 PM. Assume a program schedule on ABC of News from 6:30:00 PM to 6:59:59 PM, followed by Action Show from 7:00:00 PM to 7:59:59 PM, followed by Sports Show from 8:00:00 PM to 8:29:59 PM. Using this sample data set, we see that the linear tuning event from Device **100** can be sessionized as follows:

[0160] i. Session 1: Channel ABC from 6:55:00 PM to 6:59:59 PM tuned to Program 'News'.

[0161] ii. Session 2: Channel ABC from 7:00:00 PM to 7:59:59 PM tuned to Program 'Action Show'.

[0162] iii. Session 3: Channel ABC from 8:00:00 PM to 8:07:59 PM tuned to Program 'Sports Show.'

[0163] This is defined in more detail in the specification that Cable Television Laboratories, Inc. has published as "Audience Data Measurement Specification" referred to previously.

[0164] The same video content may air multiple times, so when viewing is being measured it is necessary to track the Program information, Program Airing information, and Channel information along with other fields.

[0165] Digital video recording viewing activity (DVR)—DVR viewing activity may be extracted from the Cable Television Laboratories, Inc. MMDM or from any source that is able to provide the data in a format suitable for this process. As a non-limiting example, a PROGRAM may be a baseball game and a PROGRAM\_AIRING may be the initial airing of the game followed by a replay later that day (thus two airings). Because a DVR recording is a recording of a video asset that airs on a certain channel at a certain time, DVR assets are also linked with Channel.

[0166] When the viewer records a program, they may record any of several airings. Also the program may air on different channels for the original and the replay. For these reasons, the DVR recording and subsequent playback must identify the Program information, Program Airing information, and Channel information along with other fields.

[0167] Video on demand viewing activity (VOD)—VOD viewing activity is extracted from the Cable Television Laboratories, Inc. MMDM or from any source that is able to provide the data in a format suitable for this process. As a non-limiting example, a program may be any VOD asset, a movie, a sporting event, an online class, etc. The concept of a PROGRAM\_AIRING does not apply to VOD. As to channel, because a VOD Program airs on the preassigned VOD channel, the channel information is not particularly informative in the context of VOD, so it is ignored for purposes of this Application.

**[0168]** Educational viewing activity (EDU)—EDU viewing activity is similar to VOD viewing in that a pre-recorded program is being viewed. This may be an educational video or any other type of video file. EDU viewing activity is often gathered from a web page that has been instrumented to capture this kind of activity.

**[0169]** Live viewing activity (LIV)—LIV viewing activity is any kind of live streaming video activity. LIV viewing activity is often gathered from a web page or cable TV network or satellite TV network that has been instrumented to capture this kind of activity.

**[0170]** Viewing types LVA, DVR, VOD, and EDU all share the characteristic that for measurement purposes, the measuring activity must be based on position in the content rather than a time element such as UTC time or local time. Measuring viewing activity based on position in the content enables creation of metrics having the same basis across LVA, DVR, VOD, and EDU viewing. Furthermore, when measuring based on position in the content, it is necessary to use a fixed reference point such as the actual start of the content. To illustrate, for a DVR recording, not every viewer will record the content beginning at the start of the program. So using position in the recording for measurement would not yield comparable results. Instead it is required to use a fixed reference point (the start of the content), in the actual content, not the start of the recording, and measure viewing from that fixed reference point. Using a fixed reference point enables consistent measurement regardless of whether the user begins recording at the beginning or some other point in the content. The same concept is applicable to LVA and VOD and EDU viewing. This is why position related fields are included in FIG. 5.

**[0171]** Resuming with the review of FIG. 2, those with ordinary skill in the art will readily recognize how to create the Video Viewing Activity Data File **130** (FIG. 5). Note that there is a single record layout for all the record types. Fields which are not used may be left blank.

**[0172]** Once the Video Viewing Activity Data File **130** has been written to the distributed file system it is ready for use by downstream processes such as my Translation Processors.

**[0173]** Depending on the criteria which the analyst intends to use in aggregating the data, various fields can be omitted

from the Video Viewing Activity Data File **130** file because the MapReduce process will not use them. I have included these additional fields to provide a comprehensive picture recognizing that one can always drop the fields that they choose not to use.

**[0174]** FIG. 3 illustrates an exemplary process for enriching the Video Viewing Activity Data File **130** with channel information. I present two alternatives: (a) loading the channel information translation table from a flat file into the memory of the computer running the Channel Translation Processor **150**, and (b) performing a join operation using channel information translation data from a database table.

**[0175]** The process begins with Channel Enrichment Process Overview **120**.

**[0176]** The Channel Translation Processor **150** requires several steps:

**[0177]** Step 1:

**[0178]** Load the Channel Information Translation Table as flat file **140** into the memory of the computer in a lookup table. This file can be presented to the Channel Translation Processor **150** as a csv file which is then read and loaded to an array in the memory of the computer.

**[0179]** A read loop such as the following can be used to load this table:

---

```

Move 0 to sub
Execute Read Loop until end of file
Read loop
  Read Channel Information Translation Table as flat file 140
  If end of file
    exit loop
  Else
    Add 1 to sub
    move SOURCE_CHANNEL_INFORMATION_IN 1810
      to CHAN-MAP-CHAN-INFO-SOURCE(sub) 1910
    move TARGET_CHANNEL_INFORMATION_IN 1820
      to CHAN-MAP-CHAN-INFO-TARGET(sub) 1920
  end-if
end loop

```

---

**[0180]** Step 2:

**[0181]** Process the Video Viewing Activity Data File **130** using a read loop as follows:

---

```

Execute Read Loop until end of file
Read Loop
  Read Video Viewing Activity Data File 130
  If end of file
    exit loop
  Else
    MOVE PROGRAM_INFO           1010 TO PROGRAM_INFO           1210
    MOVE PROGRAM_AIRING_INFO     1020 TO PROGRAM_AIRING_INFO     1220
    MOVE CHANNEL_INFO            1030 TO CHANNEL_INFO            1230
    MOVE HOUSE_INFO              1040 TO HOUSE_INFO              1240
    MOVE DEVICE_INFO             1050 TO DEVICE_INFO             1250
    MOVE VIEWER_INFO             1060 TO VIEWER_INFO             1260
    MOVE GEOGRAPHIC_INFO         1070 TO GEOGRAPHIC_INFO         1270
    MOVE DEMOGRAPHIC_INFO        1080 TO DEMOGRAPHIC_INFO        1280
    MOVE VIEWING_TYPE            1090 TO VIEWING_TYPE            1290
    MOVE VIEWING_DATE            1100 TO VIEWING_DATE            1300
    MOVE TUNE_IN_DATE_TIME        1102 TO TUNE_IN_DATE_TIME      1302
    MOVE TUNE_IN_SECOND_OF_DAY    1104 TO TUNE_IN_SECOND_OF_DAY  1304
    MOVE TUNE_OUT_DATE_TIME       1106 TO TUNE_OUT_DATE_TIME     1306
    MOVE TUNE_OUT_SECOND_OF_DAY   1108 TO TUNE_OUT_SECOND_OF_DAY  1308
    MOVE PLAYBACK_MODE            1110 TO PLAYBACK_MODE          1310
    MOVE PLAYBACK_BEG_POSITION    1120 TO PLAYBACK_BEG_POSITION  1320
    MOVE PLAYBACK_END_POSITION    1130 TO PLAYBACK_END_POSITION  1330
    MOVE TUNE_DURATION_SECONDS    1138 TO TUNE_DURATION_SECONDS  1338
    MOVE ACTIVITY_DURATION_SECONDS 1140 TO ACTIVITY_DURATION_SECONDS 1340
  
```

---

-continued

MOVE PRESENTATION_INFO	1150 TO PRESENTATION_INFO	1350
MOVE VIDEO_SERVER_INFO	1160 TO VIDEO_SERVER_INFO	1360
MOVE SYSTEM_HEALTH_INFO	1170 TO SYSTEM_HEALTH_INFO	1370
PERFORM DO_TARGET_CHANNEL_LOOKUP (see below)		
Write Prepared Video Viewing Activity Data File	160	
End loop		
DO_TARGET_CHANNEL_LOOKUP		
MOVE 'N' TO TARGET-CHANNEL-FOUND		
MOVE 'N' TO EXCEEDED-MAX-ROWS-IN-ARRAY		
PERFORM VARYING SUB		
FROM 1 BY 1		
UNTIL TARGET-CHANNEL-FOUND = 'Y'		
OR EXCEEDED-MAX-ROWS-IN-ARRAY = 'Y'		
IF CHANNEL_INFO 1030 = CHAN-MAP-CHAN-INFO-SOURCE(sub) 1910		
MOVE CHAN-MAP-CHAN-INFO-TARGET(sub) 1920		
TO CHANNEL_INFO_TARGET 1400		
MOVE 'Y' TO TARGET-CHANNEL-FOUND		
END-IF		
IF SUB > 600		
MOVE 'Y' TO EXCEEDED-MAX-ROWS-IN-ARRAY		
END-IF		
END-PERFORM		

[0182] Step 3:  
[0183] When the Process finishes reading the Video Viewing Activity Data File 130, proceed to Provide File to Downstream Process 210.

Alternative Embodiment

[0184] Instead of using a Lookup table as described above, the Channel Information Translation table may be loaded to a database table. In that case the DO\_TARGET\_CHANNEL\_LOOKUP process is done as follows:

SELECT TARGET_CHANNEL_INFORMATION_IN 1820
INTO CHANNEL_INFO_TARGET 1400
FROM CHANNEL INFORMATION TRANSLATION TABLE 142
WHERE SOURCE_CHANNEL_INFORMATION_IN 1810 = CHANNEL_INFO 1030

[0185] This completes FIG. 3.  
[0186] FIG. 4 illustrates an exemplary process for enriching the Video Viewing Activity Data File 130 with any variety of translation data. Whereas FIG. 3 was focused on translating Channel Information, FIG. 4 presents a more generalized solution which enables the analyst to translate, as a Non-limiting example, any of the following fields:  
[0187] (i) program information,  
[0188] (ii) channel information,  
[0189] (iii) house information,  
[0190] (iv) device information,  
[0191] (v) viewer information,  
[0192] (vi) geographic information,  
[0193] (vii) demographic information.  
[0194] I present two alternatives: (a) loading the data translation table from a flat file into the memory of the computer running the Data Translation Processor 154, and (b) performing a join operation using data translation data from a database table.  
[0195] The process begins with Generalized Enrichment Process Overview 124.

[0196] The Data Translation Processor 154 requires several steps:  
[0197] Step 1:  
[0198] Load the Data Translation Table as flat file 146 into the memory of the computer in a lookup table. This file can be presented to the Data Translation Processor 154 as a csv file which is then read and loaded to an array in the memory of the computer.  
[0199] A read loop such as the following can be used to load this table:

Move 0 to sub
Execute Read Loop until end of file
Read loop
Read Data Translation Table as flat file 146
If end of file
exit loop
Else
Add 1 to sub
move TRANSLATION_VALUE_TYPE-IN 2005
to TRANSLATION_VALUE_TYPE(sub) 2105
move SOURCE_TRANSLATION_VALUE-IN 2010
to SOURCE_TRANSLATION_VALUE(sub) 2110
move TARGET_TRANSLATION_VALUE-IN 2020

-continued

---

```

                to TARGET__TRANSLATION__VALUE(sub) 2120
            end-if
        end loop

```

---

**[0200]** Step 2:

**[0201]** Accept the various input parameters which will indicate which translations are being done. There is one parameter for each kind of translation which may be done. Each parameter contains a value of 'Y' or 'N' indicating whether or not that field will be translated by the Data Translation Processor **154**. The code is as follows:

---

```

        Accept program-information-translation-flag from
            job-run-param-program-information

```

---

-continued

---

```

        Accept channel-information-translation-flag from
            job-run-param-channel-information
        Accept house-information-translation-flag from
            job-run-param-house-information
        Accept device-information-translation-flag from
            job-run-param-device-information
        Accept viewer-information-translation-flag from
            job-run-param-viewer-information
        Accept geographic-information-translation-flag from
            job-run-param-geographic-information
        Accept demographic-information-translation-flag from
            job-run-param-demographic-information

```

---

**[0202]** Step 3:

**[0203]** Process the Video Viewing Activity Data File **130** using a read loop as follows:

---

```

Execute Read Loop until end of file
Read Loop
    Read Video Viewing Activity Data File 130
    If end of file
        exit loop
    Else
        IF program-information-translation-flag = 'Y'
            PERFORM DO__TARGET__PROGRAM__LOOKUP (see below)
        ELSE
            MOVE PROGRAM__INFO                1010 TO PROGRAM__INFO                1210
        END-IF
        MOVE PROGRAM__AIRING__INFO            1020 TO PROGRAM__AIRING__INFO            1220
        IF channel-information-translation-flag = 'Y'
            PERFORM DO__TARGET__CHANNEL__LOOKUP (see below)
        ELSE
            MOVE CHANNEL__INFO                1030 TO CHANNEL__INFO                1230
        END-IF
        IF house-information-translation-flag = 'Y'
            PERFORM DO__TARGET__HOUSE__LOOKUP (see below)
        ELSE
            MOVE HOUSE__INFO                 1040 TO HOUSE__INFO                 1240
        END-IF
        IF device-information-translation-flag = 'Y'
            PERFORM DO__TARGET__DEVICE__LOOKUP (see below)
        ELSE
            MOVE DEVICE__INFO                1050 TO DEVICE__INFO                1250
        END-IF
        IF viewer-information-translation-flag = 'Y'
            PERFORM DO__TARGET__VIEWER__LOOKUP (see below)
        ELSE
            MOVE VIEWER__INFO                1060 TO VIEWER__INFO                1260
        END-IF
        IF geographic-information-translation-flag = 'Y'
            PERFORM DO__TARGET__GEOGRAPHIC__LOOKUP (see below)
        ELSE
            MOVE GEOGRAPHIC__INFO            1070 TO GEOGRAPHIC__INFO            1270
        END-IF
        IF demographic-information-translation-flag = 'Y'
            PERFORM DO__TARGET__DEMOGRAPHIC__LOOKUP (see below)
        ELSE
            MOVE DEMOGRAPHIC__INFO           1080 TO DEMOGRAPHIC__INFO           1280
        END-IF
        MOVE VIEWING__TYPE                   1090 TO VIEWING__TYPE                   1290
        MOVE VIEWING__DATE                   1100 TO VIEWING__DATE                   1300
        MOVE TUNE__IN__DATE__TIME            1102 TO TUNE__IN__DATE__TIME            1302
        MOVE TUNE__IN__SECOND__OF__DAY        1104 TO TUNE__IN__SECOND__OF__DAY        1304
        MOVE TUNE__OUT__DATE__TIME           1106 TO TUNE__OUT__DATE__TIME           1306
        MOVE TUNE__OUT__SECOND__OF__DAY       1108 TO TUNE__OUT__SECOND__OF__DAY       1308
        MOVE PLAYBACK__MODE                  1110 TO PLAYBACK__MODE                  1310
        MOVE PLAYBACK__BEG__POSITION          1120 TO PLAYBACK__BEG__POSITION          1320
        MOVE PLAYBACK__END__POSITION          1130 TO PLAYBACK__END__POSITION          1330
        MOVE TUNE__DURATION__SECONDS          1138 TO TUNE__DURATION__SECONDS          1338
        MOVE ACTIVITY__DURATION__SECONDS     1140 TO ACTIVITY__DURATION__SECONDS     1340
        MOVE PRESENTATION__INFO              1150 TO PRESENTATION__INFO              1350
        MOVE VIDEO__SERVER__INFO             1160 TO VIDEO__SERVER__INFO             1360
        MOVE SYSTEM__HEALTH__INFO            1170 TO SYSTEM__HEALTH__INFO            1370

```

---

-continued

---

```

        Write Prepared Video Viewing Activity Data File 160
    End-if
End loop
DO_TARGET_PROGRAM_LOOKUP
    MOVE 'N' TO TARGET-PROGRAM-FOUND
    MOVE 'N' TO EXCEEDED-MAX-ROWS-IN-ARRAY
    PERFORM VARYING SUB
        FROM 1 BY 1
        UNTIL TARGET-PROGRAM-FOUND = 'Y'
            OR EXCEEDED-MAX-ROWS-IN-ARRAY = 'Y'
            IF TRANSLATION_VALUE_TYPE(sub) 2105 = 'PROGRAM'
                IF PROGRAM_INFO 1010 = SOURCE_TRANSLATION_VALUE(sub) 2110
                    MOVE TARGET_TRANSLATION_VALUE(sub) 2120
                        TO PROGRAM_INFO 1210
                    MOVE 'Y' TO TARGET-PROGRAM-FOUND
                END-IF
            END-IF
        END-IF
    IF SUB > 1000
        MOVE 'Y' TO EXCEEDED-MAX-ROWS-IN-ARRAY
        MOVE 'not available' TO PROGRAM_INFO 1210
    END-IF
END-PERFORM

```

---

#### Alternative Embodiment

**[0204]** Instead of using a Lookup table as described above, the Data Translation Table may be provided as a database table as shown by Data Translation Table as database Table 148. In that case the DO\_TARGET\_PROGRAM\_LOOKUP process is done as follows:

---

```

SELECT TARGET_TRANSLATION_VALUE-IN 2020
    INTO PROGRAM_INFO 1210
FROM DATA TRANSLATION TABLE 148
    WHERE TRANSLATION_VALUE_TYPE-IN 2005 = 'PROGRAM'
    AND SOURCE_TRANSLATION_VALUE-IN 2010 = PROGRAM_INFO 1010
DO_TARGET_CHANNEL_LOOKUP
    MOVE 'N' TO TARGET-CHANNEL-FOUND
    MOVE 'N' TO EXCEEDED-MAX-ROWS-IN-ARRAY
    PERFORM VARYING SUB
        FROM 1 BY 1
        UNTIL TARGET-CHANNEL-FOUND = 'Y'
            OR EXCEEDED-MAX-ROWS-IN-ARRAY = 'Y'
            IF TRANSLATION_VALUE_TYPE(sub) 2105 = 'CHANNEL'
                IF CHANNEL_INFO 1030 = SOURCE_TRANSLATION_VALUE(sub) 2110
                    MOVE TARGET_TRANSLATION_VALUE(sub) 2120 TO CHANNEL_INFO 1230
                    MOVE 'Y' TO TARGET-CHANNEL-FOUND
                END-IF
            END-IF
        END-IF
    IF SUB > 1000
        MOVE 'Y' TO EXCEEDED-MAX-ROWS-IN-ARRAY
        MOVE 'not available' TO CHANNEL_INFO 1230
    END-IF
END-PERFORM

```

---

#### Alternative Embodiment

**[0205]** Instead of using a Lookup table as described above, the Data Translation Table may be provided as a database table as shown by Data Translation Table as database Table 148. In that case the DO\_TARGET\_CHANNEL\_LOOKUP process is done as follows:

---

```

SELECT TARGET_TRANSLATION_VALUE-IN 2020
    INTO CHANNEL_INFO 1230
FROM DATA TRANSLATION TABLE 148
    WHERE TRANSLATION_VALUE_TYPE-IN 2005 = 'CHANNEL'
    AND SOURCE_TRANSLATION_VALUE-IN 2010 = CHANNEL_INFO 1030
DO_TARGET_HOUSE_LOOKUP
    MOVE 'N' TO TARGET-HOUSE-FOUND

```

-continued

---

```

MOVE 'N' TO EXCEEDED-MAX-ROWS-IN-ARRAY
PERFORM VARYING SUB
  FROM 1 BY 1
  UNTIL TARGET-HOUSE-FOUND = 'Y'
    OR EXCEEDED-MAX-ROWS-IN-ARRAY = 'Y'
    IF TRANSLATION__VALUE__TYPE(sub) 2105 = 'HOUSE'
      IF HOUSE__INFO 1040 = SOURCE__TRANSLATION__VALUE(sub) 2110
        MOVE TARGET__TRANSLATION__VALUE(sub) 2120 TO HOUSE__INFO 1240
        MOVE 'Y' TO TARGET-HOUSE-FOUND
      END-IF
    END-IF
  END-IF
  IF SUB > 1000
    MOVE 'Y' TO EXCEEDED-MAX-ROWS-IN-ARRAY
    MOVE 'not available' TO HOUSE__INFO 1240
  END-IF
END-PERFORM

```

---

#### Alternative Embodiment

**[0206]** Instead of using a Lookup table as described above, the Data Translation Table may be provided as a database table as shown by Data Translation Table as database Table 148. In that case the DO\_TARGET\_HOUSE\_LOOKUP process is done as follows:

---

```

SELECT TARGET__TRANSLATION__VALUE-IN 2020
  INTO HOUSE__INFO      1240
FROM DATA TRANSLATION TABLE 148
  WHERE TRANSLATION__VALUE__TYPE-IN 2005 = 'HOUSE'
  AND SOURCE__TRANSLATION__VALUE-IN 2010 = HOUSE__INFO      1040
DO_TARGET_DEVICE_LOOKUP.
MOVE 'N' TO TARGET-DEVICE-FOUND
MOVE 'N' TO EXCEEDED-MAX-ROWS-IN-ARRAY
PERFORM VARYING SUB
  FROM 1 BY 1
  UNTIL TARGET-DEVICE-FOUND = 'Y'
    OR EXCEEDED-MAX-ROWS-IN-ARRAY = 'Y'
    IF TRANSLATION__VALUE__TYPE(sub) 2105 = 'DEVICE'
      IF DEVICE__INFO 1050 = SOURCE__TRANSLATION__VALUE(sub) 2110
        MOVE TARGET__TRANSLATION__VALUE(sub) 2120 TO DEVICE__INFO 1250
        MOVE 'Y' TO TARGET-DEVICE-FOUND
      END-IF
    END-IF
  END-IF
  IF SUB > 1000
    MOVE 'Y' TO EXCEEDED-MAX-ROWS-IN-ARRAY
    MOVE 'not available' TO DEVICE__INFO      1250
  END-IF
END-PERFORM

```

---

#### Alternative Embodiment

**[0207]** Instead of using a Lookup table as described above, the Data Translation Table may be provided as a database table as shown by Data Translation Table as database Table 148. In that case the DO\_TARGET\_DEVICE\_LOOKUP process is done as follows:

---

```

SELECT TARGET__TRANSLATION__VALUE-IN 2020
  INTO DEVICE__INFO      1250
FROM DATA TRANSLATION TABLE 148
  WHERE TRANSLATION__VALUE__TYPE-IN 2005 = 'DEVICE'
  AND SOURCE__TRANSLATION__VALUE-IN 2010 = DEVICE__INFO      1050
DO_TARGET_VIEWER_LOOKUP.
MOVE 'N' TO TARGET-VIEWER-FOUND
MOVE 'N' TO EXCEEDED-MAX-ROWS-IN-ARRAY
PERFORM VARYING SUB
  FROM 1 BY 1
  UNTIL TARGET-VIEWER-FOUND = 'Y'
    OR EXCEEDED-MAX-ROWS-IN-ARRAY = 'Y'
    IF TRANSLATION__VALUE__TYPE(sub) 2105 = 'VIEWER'

```

-continued

---

```

        IF VIEWER_INFO 1060 = SOURCE_TRANSLATION_VALUE(sub) 2110
            MOVE TARGET_TRANSLATION_VALUE(sub) 2120 TO VIEWER_INFO 1260
            MOVE 'Y' TO TARGET-VIEWER-FOUND
        END-IF
    END-IF
    IF SUB > 1000
        MOVE 'Y' TO EXCEEDED-MAX-ROWS-IN-ARRAY
        MOVE 'not available' TO VIEWER_INFO 1260
    END-IF
END-PERFORM

```

---

#### Alternative Embodiment

**[0208]** Instead of using a Lookup table as described above, the Data Translation Table may be provided as a database table as shown by Data Translation Table as database Table **148**. In that case the DO\_TARGET\_VIEWER\_LOOKUP process is done as follows:

---

```

SELECT TARGET_TRANSLATION_VALUE-IN 2020
    INTO VIEWER_INFO 1260
FROM DATA TRANSLATION TABLE 148
    WHERE TRANSLATION_VALUE_TYPE-IN 2005 = 'VIEWER'
    AND SOURCE_TRANSLATION_VALUE-IN 2010 = VIEWER_INFO 1060
DO_TARGET_GEOGRAPHIC_LOOKUP.
MOVE 'N' TO TARGET-GEOGRAPHIC-FOUND
MOVE 'N' TO EXCEEDED-MAX-ROWS-IN-ARRAY
PERFORM VARYING SUB
    FROM 1 BY 1
    UNTIL TARGET-GEOGRAPHIC-FOUND = 'Y'
        OR EXCEEDED-MAX-ROWS-IN-ARRAY = 'Y'
        IF TRANSLATION_VALUE_TYPE(sub) 2105 = 'GEOGRAPHIC'
            IF GEOGRAPHIC_INFO 1070 = SOURCE_TRANSLATION_VALUE(sub) 2110
                MOVE TARGET_TRANSLATION_VALUE(sub) 2120 TO GEOGRAPHIC_INFO 1270
                MOVE 'Y' TO TARGET-GEOGRAPHIC-FOUND
            END-IF
        END-IF
    END-IF
    IF SUB > 1000
        MOVE 'Y' TO EXCEEDED-MAX-ROWS-IN-ARRAY
        MOVE 'not available' TO GEOGRAPHIC_INFO 1270
    END-IF
END-PERFORM

```

---

#### Alternative Embodiment

**[0209]** Instead of using a Lookup table as described above, the Data Translation Table may be provided as a database table as shown by Data Translation Table as database Table **148**. In that case the DO\_TARGET\_GEOGRAPHIC\_LOOKUP process is done as follows:

---

```

SELECT TARGET_TRANSLATION_VALUE-IN 2020
    INTO GEOGRAPHIC_INFO 1270
FROM DATA TRANSLATION TABLE 148
    WHERE TRANSLATION_VALUE_TYPE-IN 2005 = 'GEOGRAPHIC'
    AND SOURCE_TRANSLATION_VALUE-IN 2010 = GEOGRAPHIC_INFO 1070
DO_TARGET_DEMOGRAPHIC_LOOKUP.
MOVE 'N' TO TARGET-DEMOGRAPHIC-FOUND
MOVE 'N' TO EXCEEDED-MAX-ROWS-IN-ARRAY
PERFORM VARYING SUB
    FROM 1 BY 1
    UNTIL TARGET-DEMOGRAPHIC-FOUND = 'Y'
        OR EXCEEDED-MAX-ROWS-IN-ARRAY = 'Y'
        IF TRANSLATION_VALUE_TYPE(sub) 2105 = 'DEMOGRAPHIC'
            IF DEMOGRAPHIC_INFO 1080 = SOURCE_TRANSLATION_VALUE(sub) 2110
                MOVE TARGET_TRANSLATION_VALUE(sub) 2120 TO DEMOGRAPHIC_INFO 1280
                MOVE 'Y' TO TARGET-DEMOGRAPHIC-FOUND
            END-IF
        END-IF
    END-IF
    IF SUB > 1000

```

---

-continued

---

```

MOVE 'Y' TO EXCEEDED-MAX-ROWS-IN-ARRAY
MOVE 'not available' TO DEMOGRAPHIC__INFO 1280
END-IF
END-PERFORM

```

---

#### Alternative Embodiment

**[0210]** Instead of using a Lookup table as described above, the Data Translation Table may be provided as a database table as shown by Data Translation Table as database Table **148**. In that case the DO\_TARGET\_DEMOGRAPHIC\_LOOKUP process is done as follows:

---

```

SELECT TARGET__TRANSLATION__VALUE-IN 2020
      INTO DEMOGRAPHIC__INFO      1280
FROM DATA TRANSLATION TABLE 148
WHERE TRANSLATION__VALUE__TYPE-IN 2005 = 'DEMOGRAPHIC'
AND SOURCE__TRANSLATION__VALUE-IN 2010 = DEMOGRAPHIC__INFO 1080

```

---

**[0211]** Step 4:

**[0212]** When the Data Translation Processor **154** finishes reading the Video Viewing Activity Data File **130** and enriching each record as needed, proceed to Provide File to Downstream Processes **214**.

**[0213]** Note: In each case the Prepared Video Viewing Activity Data File **160** records can be written directly to a distributed file system such as, but not limited to, the Hadoop Distributed File System (HDFS) so that the prepared video viewing activity records are ready for use by downstream processes.

**[0214]** For each of these embodiments, at the completion of Data Translation Processor **154**, one record has been written to the Prepared Video Viewing Activity Data File **160** for each record in the input file. In FIG. **8** below I will review various Before and After images to provide examples of inputs and subsequent outputs from the Data Translation Processor **154**.

**[0215]** Those skilled in the art will readily recognize that the Data Translation Processor **154** is suitable for running in parallel on multiple computers simultaneously with each process creating Prepared Video Viewing Activity Data File records that can be fed into the downstream processes.

**[0216]** Note:

**[0217]** The Video Viewing Activity Data File **130** can be provided to the Data Translation Processor **154** process in any computer readable format including, but not limited to, database tables, flat files, JSON messages, and XML messages. Alternatively, such video viewing events can be collected directly from the source without the need for a Media Measurement Database **100**. In such a case, those events can still be provided as video viewing activity in a format similar to that shown in FIG. **5** for use by the Data Translation Processor **154**.

**[0218]** This concludes discussion on FIG. **4**.

**[0219]** FIG. **5** illustrates an exemplary record layout for a Video Viewing Activity Data File **130** record formatted for use as input to either the Channel Translation Processor **150** or the Data Translation Processor **154**, according to one embodiment. There is Summary Information followed by the Data Structure including field definitions.

**[0220]** FIG. **6** illustrates an exemplary record layout for a Channel Information Translation Table **140** or **142** record formatted for use as input to the Channel Translation Proces-

sor **150**, according to one embodiment. There is Summary Information followed by the Data Structure including field definitions. After the Data Structure there is a set of Sample Data.

**[0221]** FIG. **6** Example 1 shows an example of the data that would be used to translate Standard Definition and High Definition Call Signs to the corresponding Common Call Sign, according to one embodiment. This kind of translation can be used in relation to audience viewership measurement where it is desirable to combine the viewing of standard definition and high definition channels into a single call sign for the applicable channel.

**[0222]** FIG. **6** Example 2 shows an example of the data that would be used to translate selected Standard Definition and High Definition Call Signs to a Common Call Sign, and then translate all the other channel call signs to a common value, according to one embodiment. This kind of translation can be used in relation to audience viewership measurement where it is desirable to combine the viewing of standard definition and high definition channels into a single call sign for the applicable channel with the objective of simply measuring viewing of that specific channel, either by itself or as a percentage of the overall viewing that is occurring during the measurement window. By reducing the number of target channels to a single value for all of the channels that are not of interest in the analysis, the aggregation run time can be reduced substantially and the computer resource usage can be reduced as well.

**[0223]** Those skilled in the art will be able to create various combinations of the mappings to meet any number of analytical needs.

**[0224]** FIG. **7** illustrates an exemplary data structure for a Channel Mapping Definition Array Data Structure which may be used by the Channel Translation Processor **150**, according to one embodiment. There is Summary Information followed by the Data Structure including field definitions.

**[0225]** The process for loading the file in FIG. **6** to the Array in FIG. **7** is described in FIG. **3** Step 1.

**[0226]** FIG. **8** illustrates an exemplary record layout for a Data Translation Table **146** or **148** record formatted for use as input to the Data Translation Processor **154**, according to one embodiment. There is Summary Information followed by the



Data Structure including field definitions. After the Data Structure there is a set of Sample Data.

**[0227]** FIG. 8 Example 1 relates to translating PROGRAM information, according to one embodiment. This figure shows an example of the data that would be used to translate program type. This non-limiting example shows keeping the type of “SPORT” unchanged while replacing the other values with a value of “OTHER”. This kind of translation can be used in relation to audience viewership measurement where the analyst needs to compare viewing of one kind of program with all the other viewing that is happening concurrently.

**[0228]** FIG. 8 Example 2 shows an example of the data that would be used to translate Standard Definition and High Definition Call Signs to a Common Call Sign, according to one embodiment.

**[0229]** This kind of translation can be used in relation to audience viewership measurement where it is desirable to combine the viewing of standard definition and high definition channels into a single call sign for the applicable channel.

**[0230]** FIG. 8 Example 3 shows an example of the data that would be used to translate selected Standard Definition and High Definition Call Signs to a Common Call Sign, and then translate all the other channel call signs to a common value, according to one embodiment. This kind of translation can be used in relation to audience viewership measurement where it is desirable to combine the viewing of standard definition and high definition channels into a single call sign for the applicable channel with the objective of simply measuring that specific channel. By reducing the number of target channels to a single value for all of the channels that are not of interest in the analysis, the aggregation run time can be reduced substantially and the computer resources can be reduced as well.

**[0231]** FIG. 8 Example 4 relates to translating HOUSE information, according to one embodiment. This figure shows an example of the data that would be used to translate house type. This non-limiting example shows keeping the type of “HOUSE” unchanged while replacing the other values (APT, OFFICE, SCHOOL) with a value of “OTHER”. This kind of translation can be used in relation to audience viewership measurement where the analyst needs to compare viewing of one kind of house with all the other viewing that is happening concurrently (HOUSE viewing vs. combined viewing by APT, OFFICE, and SCHOOL).

**[0232]** FIG. 8 Example 5 relates to translating DEVICE information, according to one embodiment. This figure shows an example of the data that would be used to translate device type. This non-limiting example shows keeping the type of “STB” unchanged while replacing the other values (IPTV, TABLET, PHONE) with a value of “OTHER”. This kind of translation can be used in relation to audience viewership measurement where the analyst needs to compare viewing of one kind of device with all the other viewing that is happening concurrently (STB viewing vs. combined viewing by IPTV, TABLET, and PHONE).

**[0233]** FIG. 8 Example 6 relates to translating VIEWER information, according to one embodiment. This figure shows an example of the data that would be used to translate viewer type. This non-limiting example shows keeping the type of “PARENT” unchanged while replacing the other values (CHILD, STUDENT) with a value of “OTHER”. This kind of translation can be used in relation to audience viewership measurement where the analyst needs to compare viewing of

one kind of viewer with all the other viewing that is happening concurrently (PARENT viewing vs. combined viewing by CHILD and STUDENT).

**[0234]** FIG. 8 Example 7 relates to translating GEOGRAPHIC information, according to one embodiment. This figure shows an example of the data that would be used to translate geographic information. This non-limiting example shows translating from a more detailed view to a higher view. Thus each of the detailed zip codes values is replaced by a value with only the first three digits of the zip code followed by XX. This kind of translation can be used in relation to audience viewership measurement where the analyst needs to aggregate the viewing of the individual zip codes to a higher level code.

**[0235]** FIG. 8 Example 8 relates to translating DEMOGRAPHIC information, according to one embodiment. This figure shows an example of the data that would be used to translate demographic information. This non-limiting example shows translating from a more detailed view which lists age ranges of children and adults to a higher view which simply lists CHILD and ADULT. Thus each of the detailed age range values for children is replaced by the value “CHILD” and the adult age ranges are replaced by “ADULT”. This kind of translation can be used in relation to audience viewership measurement where the analyst needs to aggregate the viewing of the various age ranges to a more general value.

**[0236]** FIG. 8 Summary

**[0237]** Each of the above non-limiting examples shows translating a more detailed value to a less detailed value. By reducing the number of distinct values to be used in the aggregation process, the aggregation run time can be reduced substantially and the computer resource usage can be reduced as well.

**[0238]** Those skilled in the art will be able to create various combinations of the mappings to meet any number of analytical needs.

**[0239]** FIG. 9 illustrates an exemplary data structure for a Generalized Mapping Definition Array Data Structure which may be used by the Data Translation Processor 154, according to one embodiment. There is Summary Information followed by the Data Structure including field definitions.

**[0240]** The process for loading the file in FIG. 8 to the Array in FIG. 9 is described in FIG. 4 Step 1.

**[0241]** FIG. 10 illustrates various non-limiting examples which provide ‘Before’ and ‘After’ images of Linear Tuning Activity (LTA) records to illustrate what the Channel Translation Processor 150 and Data Translation Processor 154 do, according to one embodiment.

**[0242]** FIG. 10 Example 1 relates to translating GEOGRAPHIC information, according to one embodiment. The “Before” shows Video Viewing Activity Data File 130 records with detailed Geographic Id’s (Zip codes)—reference field 1070.

**[0243]** The “After” shows Prepared Video Viewing Activity Data File 160 records with detailed Geographic Id’s (Zip codes) replaced by a higher level code—reference field 1270.

**[0244]** FIG. 10 Example 2 relates to translating CHANNEL information, according to one embodiment. The “Before” shows Video Viewing Activity Data File 130 records with both Standard Definition and High Definition Call Signs—reference field 1030.

[0245] The “After” shows Prepared Video Viewing Activity Data File 160 records with detailed call signs replaced by Common Call Signs—reference field 1230.

[0246] FIG. 10 Example 3 relates to translating CHANNEL information, according to one embodiment. The “Before” shows Video Program Viewing Activity Data File 130 records with both Standard Definition and High Definition Call Signs and other call signs for the various channels—reference field 1030.

[0247] The “After” shows Prepared Video Program Viewing Activity Data File 160 records with the Standard Definition and High Definition Call Signs mapped to a common value for the channel of interest (HIST) and the other call signs mapped to “OTHER”—reference field 1230.

[0248] FIG. 10 Example 4 relates to translating VIEWING DEVICE information, according to one embodiment. The “Before” shows Video Viewing Activity Data File 130 records with various kinds of viewing devices (STB, IPTV, TABLET, PHONE)—reference field 1050.

[0249] The “After” shows Prepared Video Viewing Activity Data File 160 records with the device type of “STB” unchanged while IPTV, TABLET, PHONE have been mapped to “OTHER”—reference field 1250.

[0250] FIG. 10 Example 5 relates to translating HOUSE information, according to one embodiment. The “Before” shows Video Viewing Activity Data File 130 records with various kinds of house information values (HOUSE, APT, OFFICE, SCHOOL)—reference field 1040.

[0251] The “After” shows Prepared Video Viewing Activity Data File 160 records with the house information type of “HOUSE” unchanged while APT, OFFICE, and SCHOOL have been mapped to “OTHER”—reference field 1240.

[0252] FIG. 10 Example 6 relates to translating VIEWER information, according to one embodiment. The “Before” shows Video Viewing Activity Data File 130 records with various kinds of viewer information values (PARENT, CHILD, STUDENT)—reference field 1060.

[0253] The “After” shows Prepared Video Viewing Activity Data File 160 records with the viewer information type of “PARENT” unchanged while CHILD and STUDENT have been mapped to “OTHER”—reference field 1260.

[0254] FIG. 10 Example 7 relates to translating DEMOGRAPHIC information, according to one embodiment. The “Before” shows Video Viewing Activity Data File 130 records with various kinds of demographic values (1-3 YR, 4-5 YR, 6-8 YR, 21-44 YR, 45-54 YR)—reference field 1080.

[0255] The “After” shows Prepared Video Viewing Activity Data File 160 records with the demographic values mapped to new, summary values (1-3 YR, 4-5 YR, 6-8 YR mapped to “CHILD” and 21-44 YR, 45-54 YR mapped to “PARENT”)—reference field 1280.

[0256] FIG. 11 illustrates various non-limiting examples which provide ‘Before’ and ‘After’ images of Linear Viewing Activity (LVA), Digital Video Recorder viewing activity (DVR), Video On Demand viewing activity (VOD), Educational viewing (EDU), and Live viewing activity (LIV) records to illustrate what the Channel Translation Processor 150 and Data Translation Processor 154 do, according to one embodiment.

[0257] FIG. 11 Example 1 relates to translating PROGRAM information, according to one embodiment. The “Before” shows Video Viewing Activity Data File 130 records with a variety of Program information values (SPORT, DOCU, NATU, NEWS)—reference field 1010.

[0258] The “After” shows Prepared Video Viewing Activity Data File 160 records with only two values (SPORT and OTHER) with the values of DOCU, NATU, NEWS having been replaced by OTHER—reference field 1210. Thus a more detailed value has been replaced by a summary value.

[0259] FIG. 11 Example 2 relates to translating CHANNEL information, according to one embodiment. The “Before” shows Video Viewing Activity Data File 130 records with both Standard Definition and High Definition Call Signs—reference field 1030.

[0260] The “After” shows Prepared Video Viewing Activity Data File 160 records with detailed call signs replaced by Common Call Signs—reference field 1230.

[0261] FIG. 11 Example 3 relates to translating CHANNEL information, according to one embodiment. The “Before” shows Video Viewing Activity Data File 130 records with both Standard Definition and High Definition Call Signs and other call signs for the various channels—reference field 1030.

[0262] The “After” shows Prepared Video Viewing Activity Data File 160 records with the Standard Definition and High Definition Call Signs mapped to a common value for the channel of interest (HIST) and the other call signs mapped to “OTHER”—reference field 1230.

[0263] FIG. 11 Example 4 relates to translating HOUSE information, according to one embodiment. The “Before” shows Video Viewing Activity Data File 130 records with various kinds of house information values (HOUSE, APT, OFFICE, SCHOOL)—reference field 1040.

[0264] The “After” shows Prepared Video Viewing Activity Data File 160 records with the house information type of “HOUSE” unchanged while APT, OFFICE, and SCHOOL have been mapped to “OTHER”—reference field 1240.

[0265] FIG. 11 Example 5 relates to translating VIEWING DEVICE information, according to one embodiment. The “Before” shows Video Viewing Activity Data File 130 records with various kinds of viewing devices (STB, IPTV, TABLET, PHONE)—reference field 1050.

[0266] The “After” shows Prepared Video Viewing Activity Data File 160 records with the device type of “STB” unchanged while IPTV, TABLET, PHONE have been mapped to “OTHER”—reference field 1250.

[0267] FIG. 11 Example 6 relates to translating VIEWER information, according to one embodiment. The “Before” shows Video Viewing Activity Data File 130 records with various kinds of viewer information values (PARENT, CHILD, STUDENT)—reference field 1060.

[0268] The “After” shows Prepared Video Viewing Activity Data File 160 records with the viewer information type of “PARENT” unchanged while CHILD and STUDENT have been mapped to “OTHER”—reference field 1260.

[0269] FIG. 11 Example 7 relates to translating GEOGRAPHIC information, according to one embodiment. The “Before” shows Video Viewing Activity Data File 130 records with detailed Geographic Id’s (Zip codes)—reference field 1070.

[0270] The “After” shows Prepared Video Viewing Activity Data File 160 records with detailed Geographic Id’s (Zip codes) replaced by a higher level code—reference field 1270.

[0271] FIG. 11 Example 8 relates to translating DEMOGRAPHIC information, according to one embodiment. The “Before” shows Video Viewing Activity Data File 130 records with various kinds of demographic values (1-3 YR, 4-5 YR, 6-8 YR, 21-44 YR, 45-54 YR)—reference field 1080.

[0272] The “After” shows Prepared Video Viewing Activity Data File 160 records with the demographic values mapped to new, summary values (1-3 YR, 4-5 YR, 6-8 YR mapped to “CHILD” and 21-44 YR, 45-54 YR mapped to “PAR-ENT”)—reference field 1280.

[0273] FIG. 11 Summary

[0274] Each of the above non-limiting examples shows translating a more detailed value to a less detailed value. By reducing the number of distinct values to be used in the aggregation process, the aggregation run time can be reduced substantially and the computer resource usage can be reduced as well.

[0275] FIG. 12 illustrates an exemplary record layout for a Prepared Video Viewing Activity Data File 160 record created by either the Channel Translation Processor 150 or the Data Translation Processor 154, according to one embodiment. There is Summary Information followed by the Data Structure including field definitions.

#### Alternative Embodiments

[0276] Although the description above contains much specificity, this should not be construed as limiting the scope of the embodiments but as merely providing illustrations of some of several embodiments. As a nonlimiting example, additional qualifiers may be added along with those provided.

[0277] Scope of Viewer Interaction Data Loaded

[0278] I presently contemplate that the Channel Translation Processor 150 and Data Translation Processor 154 will each process viewer interaction data for whatever set of viewing activity is provided to it. This may be one Video Program at a time, one hour of the day, a primetime television viewing period, an entire 24 hour day of viewing, a week of viewing, a weekly program schedule time slot, or another time period decided by the analyst. Another embodiment may simply process viewing activity within the context of a single program, or a single advertisement, or some other combination.

[0279] Identifiers for Data

[0280] I presently contemplate using a combination of numeric and mnemonics for the various fields such as program info, program airing info, channel info, house info, device info, viewer info, geographic info, demographic info, viewing type, video server identifiers, system health info, and other similar fields, but another embodiment could use only numeric values as identifiers with links to reference tables for the descriptions of the numeric identifiers or only mnemonic identifiers.

[0281] Programming Algorithm Scope

[0282] I presently contemplate executing the algorithms described herein separately in some sequence, but another embodiment could combine multiple simple algorithms into fewer complex algorithms.

[0283] Receiving Date and Time Information

[0284] I presently contemplate receiving all of the date and time values in local time, but another embodiment may provide these in Coordinated Universal Time (UTC time).

[0285] General Information

[0286] I presently contemplate using variables having the data types and field sizes shown, but another embodiment may use variables with different data types and field sizes to accomplish a similar result.

[0287] I presently contemplate using record layouts similar to those defined herein, but another embodiment may use a different record layout or record layouts to accomplish a similar result.

[0288] As a nonlimiting example, another embodiment may use database tables or other objects instead of record layouts similar to those I have defined herein to accomplish a similar result while still working within the spirit and scope of this disclosure.

[0289] Implementation Information

[0290] I presently contemplate using Linux operating system, but another embodiment may use a different operating system.

[0291] I presently contemplate using the COBOL language, but another embodiment may use Java or Python or some other language.

[0292] General Remarks

[0293] It will be apparent to those of ordinary skill in the art that various changes and modifications may be made which clearly fall within the scope of the embodiments revealed herein. In describing an embodiment illustrated in the drawings, specific terminology has been used for the sake of clarity. However, the embodiments are not intended to be limited to the specific terms so selected, and it is to be understood that each specific term includes all technical equivalents which operate in a similar manner to accomplish a similar purpose.

[0294] In general, it will be apparent to one of ordinary skill in the art that various embodiments described hereinabove may be implemented in many different embodiments of software, firmware, and/or hardware, or modules thereof. The software code or specialized control hardware used to implement some of the present embodiments is not limiting of the present embodiment. For example, the embodiments described hereinabove may be implemented in computer software using any suitable computer software language type such as, for example, Python or JAVA or COBOL using, for example, conventional or object-oriented techniques. Such software may be stored on any type of suitable computer-readable medium or media such as, for example, a magnetic or optical storage medium. Thus, the operation and behavior of the embodiments are described in COBOL language purely as a matter of convenience. It is clearly understood that artisans of ordinary skill would be able to design software and control hardware to implement the embodiments presented in the language of their choice based on the description herein with only a reasonable effort and without undue experimentation.

[0295] The processes associated with the present embodiments may be executed by programmable equipment, such as computers. Software or other sets of instructions that may be employed to cause programmable equipment to execute the processes may be stored in any storage device, such as, for example, a computer system (non-volatile) memory, a compact disk, an optical disk, magnetic tape, or magnetic disk. Furthermore, some of the processes may be programmed when the computer system is manufactured or via a computer-readable medium.

[0296] It can also be appreciated that certain process aspects disclosed herein may be performed using instructions stored on a computer-readable memory medium or media that direct a computer or computer system to perform process steps. A computer-readable medium may include, for example, memory devices such as diskettes, compact discs of both read-only and read/write varieties, optical disk drives, memory sticks, and hard disk drives. A computer-readable medium may also include memory storage that may be physical, virtual, permanent, temporary, semi-permanent and/or semi-temporary.

[0297] In various embodiments disclosed herein, a single component or algorithm may be replaced by multiple components or algorithms, and multiple components or algorithms may be replaced by a single component or algorithm, to perform a given function or functions. Except where such substitution would not be operative to implement the embodiments disclosed herein, such substitution is within the scope presented herein. Thus any element expressed herein as a means or a method for performing a specified function is intended to encompass any way of performing that function including, for example, a combination of elements that performs that function. Therefore, any means or method that can provide such functionalities may be considered equivalents to the means or methods shown herein.

[0298] It can be appreciated that the “data analysis computer system” may be, for example, any computer system capable of running the Translator Processors described herein, whether it be a one node system or a system with thousands of nodes. In an alternative embodiment, it may be a relational database server.

[0299] While various embodiments have been described herein, it should be apparent, however, that various modifications, alterations and adaptations to those embodiments may occur to persons skilled in the art with the attainment of some or all of the advantages described herein. The disclosed embodiments are therefore intended to include all such modifications, alterations and adaptations without departing from the scope and spirit of the embodiments presented herein as set forth in the appended claims.

[0300] Accordingly, the scope should be determined not by the embodiments illustrated, but by the appended claims and their legal equivalents.

#### CONCLUSIONS, RAMIFICATIONS, AND SCOPE

[0301] From the description above, a number of advantages of some embodiments of my Channel Translation Processor 150 and Data Translation Processor 154 and its supporting processes become evident:

[0302] In this specification I have taught how to reduce the workload on the Hadoop MapReduce framework by translating various values from detailed values to summary values prior to sending the data files to the downstream processes. By implementing the teachings described in this specification, an analyst can reduce the number of unique keys going into the Reduce part of the MapReduce process by 2, 5, 10, and even 100 times. This huge reduction in the unique keys results in a significant reduction in run time and computing resources needed to run the analytical study. This allows an analyst to get answers faster and to run additional analytical studies with the same or less computer hardware.

[0303] Additionally, by teaching how to translate multiple kinds of values (program information, channel information, house information, device information, viewer information, geographic information, demographic information) in a single program run, the analyst can create studies which combine multiple dimensions in one run thus being able to slice-and-dice the data in numerous ways to understand how different content is consumed. This provides a framework for creating business value through in-depth analytics.

[0304] Also, by implementing my teaching, an analyst can avoid additional expensive database extracts which may otherwise be needed to create the alternative versions of the video viewing activity files which could be fed into downstream processes.

[0305] Once the data translations are applied, the resulting prepared file is ready to be used by downstream processes.

[0306] This method of translating various detailed values to summary values prior to feeding the Video Viewing Activity Data files into the downstream processes is a novel technique that has not been taught previously. Using this technique I am able to analyze larger data sets using less hardware than was possible previously.

#### SUMMARY

[0307] In accordance with one embodiment, I have disclosed a computer-implemented method of using Linear, DVR, VOD, and streaming video viewing activity data as input to a data translation processor which prepares that video viewing activity for more efficient downstream processing by translating detailed values to aggregated values according to analyst defined translation rules in preparation for ingestion by a MapReduce Framework with the result that the MapReduce Framework needs to process less data in order to create analytical studies of second-by-second viewing activity for program, channel, house, device, viewer, demographic, and geographic attributes. The source data may be extracted from a database defined according to the Cable Television Laboratories, Inc. Media Measurement Data Model defined in “Audience Data Measurement Specification” as “OpenCable™ Specifications, Audience Measurement, Audience Measurement Data Specification” document OC-SP-AMD-101-130502 or any similar format. An analyst can use Hadoop to run more studies in less time with less hardware thus gaining greater insights into viewing activity at lower cost.

1-20. (canceled)

21. A method, comprising:

receiving, by a computing system, video-viewing-activity data comprising a plurality of entries; and

generating, by the computing system, for each entry of the plurality of entries, and responsive to a determination by the computing system that a corresponding entry in a data-translation table comprises a substitute value for the entry, a plurality of records representing a time range over which a video-asset-viewing device displayed a video asset associated with the entry, each record of the plurality of records corresponding to an interval of the time range, and each record of the plurality of records comprising the substitute value.

22. The method of claim 21, further comprising:

grouping, by the computing system and for each entry of the plurality of entries, the plurality of records with a group of records associated with the video-viewing-activity data; and

determining, by the computing system and based on a portion of the group, a video-viewing-activity metric for the video-viewing-activity data.

23. The method of claim 22, further comprising determining, by the computing system and for each record of the portion, that the record corresponds to one or more time criteria for the video-viewing-activity metric, and wherein determining the video-viewing-activity metric comprises determining a number of records in the portion.

24. The method of claim 23, further comprising determining, by the computing system, for each record in a subset of the portion, and based on its substitute value, that the record corresponds to one or more criteria, for the video-viewing-activity metric, other than time, and wherein determining the

video-viewing-activity metric comprises determining a ratio of a number of records in the subset to the number of records in the portion.

**25.** The method of claim **21**, further comprising:

receiving, by the computing system, one or more user-defined parameters for the data-translation table, the one or more user-defined parameters comprising, for each entry of the plurality of entries, a value corresponding to the entry and the substitute value; and

generating, by the computing system and based on the one or more user-defined parameters, the data-translation table.

**26.** The method of claim **21**, wherein the data-translation table comprises multiple entries comprising a common substitute value, each entry of the multiple entries comprising the common substitute value and a value different from each other entry of the multiple entries, the method further comprising:

determining, by the computing system, that a first entry of the plurality of entries corresponds to an entry of the multiple entries; and

determining, by the computing system, that a second entry of the plurality of entries corresponds to a different entry of the multiple entries.

**27.** The method of claim **21**, further comprising determining, by the computing system and for each entry of the plurality of entries, that the entry comprises data corresponding to one or more video-content-description criteria of the corresponding entry.

**28.** The method of claim **21**, further comprising determining, by the computing system and for each entry of the plurality of entries, that the entry comprises data corresponding to one or more viewer-geographic-location criteria of the corresponding entry.

**29.** The method of claim **21**, further comprising determining, by the computing system and for each entry of the plurality of entries, that the entry comprises data corresponding to one or more viewer-demographic criteria of the corresponding entry.

**30.** The method of claim **21**, further comprising determining, by the computing system and for each entry of the plurality of entries, that the entry comprises data corresponding to one or more video-asset-viewing-device-characteristic criteria of the corresponding entry.

**31.** The method of claim **21**, further comprising determining, by the computing system and for each entry of the plurality of entries, that the substitute value comprises a truncated version of a value in the entry.

**32.** The method of claim **21**, further comprising determining, by the computing system and for each entry of the plurality of entries, that the substitute value comprises a descriptor describing a value in the entry and one or more other values in other entries, of the plurality of entries, that correspond to entries, in the data-translation table, comprising the substitute value.

**33.** The method of claim **21**, further comprising determining, by the computing system and for each entry of the plurality of entries, that an element, of an array comprising the data-translation table, comprises the corresponding entry.

**34.** The method of claim **21**, further comprising performing, by the computing system, a database-table-join operation on a database table comprising the plurality of entries and a database table comprising the data-translation table.

**35.** A method, comprising:

for each video-viewing-activity event of a plurality of video-viewing-activity events:

determining, by a computing system, that the video-viewing-activity event corresponds to an entry in a data-translation table; and

generating, by the computing system, a plurality of records that each represent a different interval of a time range over which the video-viewing-activity event occurred and comprise a value from the entry.

**36.** The method of claim **35**, further comprising:

aggregating, by the computing system and for each video-viewing-activity event of the plurality of video-viewing-activity events, the plurality of records into a set of records associated with the plurality of video-viewing-activity events; and

determining, by the computing system and based on a ratio of a number of records in a subset, of the set, corresponding to one or more criteria of a video-viewing-activity metric to a number of records in the set, the video-viewing-activity metric.

**37.** The method of claim **35**, wherein determining that the video-viewing-activity event corresponds to the entry comprises determining that the video-viewing-activity event corresponds to at least one of a video-content-description criterion of the entry, a viewer-geographic-location criterion of the entry, a viewer-demographic criterion of the entry, or a video-asset-viewing-device-characteristic criterion of the entry.

**38.** A method, comprising:

determining, by a computing system and based on data generated by a plurality of different video-asset-viewing devices, a plurality of video-viewing-activity events associated with the plurality of different video-asset-viewing devices; and

determining, by the computing system, for each video-viewing-activity event of the plurality of video-viewing-activity events, and based on at least one of video content associated with the video-viewing-activity event, a geographic location associated with the video-viewing-activity event, a viewer demographic associated with the video-viewing-activity event, or a video-asset-viewing-device characteristic associated with the video-viewing-activity event, that the video-viewing-activity event corresponds to an entry in a data-translation table.

**39.** The method of claim **38**, further comprising generating, by the computing system and for each video-viewing-activity event of the plurality of video-viewing-activity events, a plurality of records that each represent a different interval of a time range over which the video-viewing-activity event occurred and comprise a value from the entry.

**40.** The method of claim **39**, further comprising:

aggregating, by the computing system and for each video-viewing-activity event of the plurality of video-viewing-activity events, the plurality of records into a set of records associated with the plurality of video-viewing-activity events; and

determining, by the computing system and based on a ratio of a number of records in a subset, of the set, corresponding to one or more criteria of a video-viewing-activity metric to a number of records in the set, the video-viewing-activity metric.

\* \* \* \* \*